

日本国特許庁
JAPAN PATENT OFFICE

18.11.03

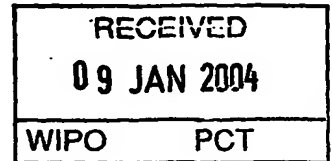
別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: 2003年 3月27日

出願番号
Application Number: 特願2003-089179
[ST. 10/C]: [JP 2003-089179]

出願人
Applicant(s): 松下電器産業株式会社

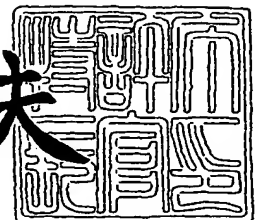


PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

2003年12月22日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



【書類名】 特許願

【整理番号】 2033850100

【あて先】 特許庁長官殿

【国際特許分類】 G10L

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

【氏名】 芳澤 伸一

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100109210

【弁理士】

【氏名又は名称】 新居 広守

【先の出願に基づく優先権主張】

【出願番号】 特願2002-338652

【出願日】 平成14年11月21日

【手数料の表示】

【予納台帳番号】 049515

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0213583

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 標準モデル作成装置および標準モデル作成方法

【特許請求の範囲】

【請求項1】 事象の出力確率によって定義される認識用のモデルである標準モデルを作成する装置であって、

特定の対象を認識するために予め作成されたモデルである1以上の参照モデルを記憶する参照モデル記憶手段と、

前記参照モデル記憶手段に記憶された1以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段と、

外部から参照モデルを取得して前記参照モデル記憶手段に格納すること、および、参照モデルを作成して前記参照モデル記憶手段に格納することの少なくとも一方を行う参照モデル準備手段と

を備えることを特徴とする標準モデル作成装置。

【請求項2】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識用のモデルである標準モデルを作成する装置であって、

特定の対象を認識するために予め作成されたモデルである1以上の参照モデルを記憶する参照モデル記憶手段と、

前記参照モデル記憶手段に記憶された1以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段と、

外部から参照モデルを取得して前記参照モデル記憶手段に格納すること、および、参照モデルを作成して前記参照モデル記憶手段に格納することの少なくとも一方を行う参照モデル準備手段と

を備えることを特徴とする標準モデル作成装置。

【請求項3】 前記標準モデル作成装置は、さらに、

認識の対象に関する情報である利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から1以上の参照モデルを選択する参照モデル選択手段とを備え、

前記標準モデル作成手段は、前記参照モデル選択手段が選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 1 又は 2 記載の標準モデル作成装置。

【請求項 4】 事象の出力確率によって定義される認識用のモデルである標準モデルを作成する装置であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルを記憶する参照モデル記憶手段と、

前記参照モデル記憶手段に記憶された 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段と、

認識の対象に関する情報である利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する参照モデル選択手段とを備え、

前記標準モデル作成手段は、前記参照モデル選択手段が選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする標準モデル作成装置。

【請求項 5】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識用のモデルである標準モデルを作成する装置であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルを記憶する参照モデル記憶手段と、

前記参照モデル記憶手段に記憶された 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段と、

認識の対象に関する情報である利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する参照モデル選択手段とを備え、

前記標準モデル作成手段は、前記参照モデル選択手段が選択した参照モデルに

対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする標準モデル作成装置。

【請求項 6】 前記標準モデル作成装置は、さらに、

認識の対象に関する情報である利用情報を作成する利用情報作成手段を備え、

前記参照モデル選択手段は、作成された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する

ことを特徴とする請求項 1～5 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 7】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

認識の対象に関する情報である利用情報を前記端末装置から受信する利用情報受信手段を備え、

前記参照モデル選択手段は、作成された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する

ことを特徴とする請求項 1～5 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 8】 前記標準モデル作成手段は、作成する標準モデルの仕様に関する情報である仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 1～7 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 9】 事象の出力確率によって定義される認識用のモデルである標準モデルを作成する装置であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルを記憶する参照モデル記憶手段と、

前記参照モデル記憶手段に記憶された 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計

算することによって標準モデルを作成する標準モデル作成手段とを備え、

前記標準モデル作成手段は、作成する標準モデルの仕様に関する情報である仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする標準モデル作成装置。

【請求項 10】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識用のモデルである標準モデルを作成する装置であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルを記憶する参照モデル記憶手段と、

前記参照モデル記憶手段に記憶された 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段とを備え、

前記標準モデル作成手段は、作成する標準モデルの仕様に関する情報である仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする標準モデル作成装置。

【請求項 11】 前記仕様情報は、標準モデルを使用するアプリケーションプログラムの種類に対応づけられた仕様を示す

ことを特徴とする請求項 8～10 のいずれか 1 項に記載の標準モデル作成装置

。

【請求項 12】 前記標準モデル作成装置は、さらに、標準モデルを使用するアプリケーションプログラムと標準モデルの仕様との対応を示すアプリケーション仕様対応データベースを前記仕様情報として保持する仕様情報保持手段を備え、

前記標準モデル作成手段は、前記仕様情報保持手段に保持されたアプリケーション仕様対応データベースから、起動されるアプリケーションプログラムに対応する仕様を読み出し、読み出した仕様に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 8～10 のいずれか 1 項に記載の標準モデル作成装置

。

【請求項 13】 作成する標準モデルの仕様に関する情報である仕様情報を作成する仕様情報作成手段を備え、

前記標準モデル作成手段は、前記仕様情報作成手段が作成した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 8～10 のいずれか 1 項に記載の標準モデル作成装置

。

【請求項 14】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

作成する標準モデルの仕様に関する情報である仕様情報を前記端末装置から受信する仕様情報受信手段を備え、

前記標準モデル作成手段は、前記仕様情報受信手段が受信した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 8～10 のいずれか 1 項に記載の標準モデル作成装置

。

【請求項 15】 前記参照モデルおよび前記標準モデルは、1 以上のガウス分布を用いて表現され、

前記標準モデル作成手段は、前記仕様情報に基づいて、前記標準モデルの構造として、少なくとも前記標準モデルの混合分布数を決定する

ことを特徴とする請求項 8～14 のいずれか 1 項に記載の標準モデル作成装置

。

【請求項 16】 前記参照モデルおよび前記標準モデルは、1 以上のガウス分布を用いて表現され、

前記参照モデル記憶手段は、少なくとも 1 対の参照モデルの混合分布数が異なる参照モデルを記憶し、

前記標準モデル作成手段は、少なくとも1対の参照モデルの混合分布数が異なる参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項1～15のいずれか1項に記載の標準モデル作成装置

。

【請求項17】 前記標準モデル作成装置は、さらに、

前記標準モデル作成手段が作成した標準モデルを記憶する標準モデル記憶手段を備える

ことを特徴とする請求項1～16のいずれか1項に記載の標準モデル作成装置

。

【請求項18】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

前記標準モデル作成手段が作成した標準モデルを前記端末装置に送信する標準モデル送信手段を備える

ことを特徴とする請求項1～17のいずれか1項に記載の標準モデル作成装置

。

【請求項19】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

前記端末装置から送信される参照モデルを受信する参照モデル受信手段を備え

、

前記標準モデル作成手段は、少なくとも前記参照モデル受信手段が受信した参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算する

ことを特徴とする請求項1～18のいずれか1項に記載の標準モデル作成装置

。

【請求項20】 前記参照モデル準備手段は、さらに、前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 1 または 2 記載の標準モデル作成装置。

【請求項 2 1】 前記参照モデル準備手段は、認識の対象に関する情報である利用情報、および作成する標準モデルの仕様に関する情報である仕様情報の少なくとも一方に基づいて、前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 2 0 記載の標準モデル作成装置。

【請求項 2 2】 前記標準モデル作成装置は、さらに、認識の対象に関する情報である利用情報、および作成する標準モデルの仕様に関する情報である仕様情報の少なくとも一方と、前記参照モデル記憶手段に記憶された参照モデルに基づいて、前記利用情報および前記仕様情報の少なくとも一方と前記参照モデルとの類似度情報を作成する類似度情報作成手段を備え、

前記参照モデル準備手段は、前記類似度情報作成手段が作成した類似度情報に基づいて、前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行うか否かを決定する

ことを特徴とする請求項 2 0 または 2 1 記載の標準モデル作成装置。

【請求項 2 3】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

前記端末装置から送信される参照モデルを受信する参照モデル受信手段を備え、

前記参照モデル準備手段は、前記参照モデル受信手段が受信した参照モデルを用いて前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 2 0 ～ 2 2 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 2 4】 前記標準モデル作成手段は、

作成する標準モデルの構造を決定する標準モデル構造決定部と、

構造が決定された前記標準モデルを特定する統計量の初期値を決定する初期標準モデル作成部と、

前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を推定して計算する統計量推定部とを有する

ことを特徴とする請求項 1～23 のいずれか 1 項に記載の標準モデル作成装置

。【請求項 25】 前記初期標準モデル作成部は、前記統計量推定部が標準モデルの統計量を計算するために用いる、1 以上の前記参照モデルを用いて前記標準モデルを特定する統計量の初期値を決定する

ことを特徴とする請求項 24 記載の標準モデル作成装置。

【請求項 26】 前記初期標準モデル作成部は、標準モデルの種類を識別するクラス ID に基づいて、前記初期値を決定する

ことを特徴とする請求項 24 記載の標準モデル作成装置。

【請求項 27】 前記初期標準モデル作成部は、前記参照モデルからクラス ID を特定し、特定したクラス ID に対応づけられた初期値を前記初期値と決定する

ことを特徴とする請求項 26 記載の標準モデル作成装置。

【請求項 28】 前記初期標準モデル作成部は、前記クラス ID と前記初期値と前記参照モデルとの対応を示す対応表を保持し、前記対応表に従って、前記初期値を決定する

ことを特徴とする請求項 27 記載の標準モデル作成装置。

【請求項 29】 前記初期標準モデル作成部は、前記クラス ID が対応づけられた初期値であるクラス ID 付き初期標準モデル、または、前記クラス ID が対応づけられた参照モデルであるクラス ID 付き参照モデルを作成または外部から取得することによって、前記対応表を生成する

ことを特徴とする請求項 28 記載の標準モデル作成装置。

【請求項 30】 前記標準モデル構造決定部は、前記標準モデルの構造を混合分布数が M_f 個 ($M_f \geq 1$) の混合ガウス分布と決定し、

前記統計量推定部は、混合ガウス分布により表現された前記標準モデル

【数1】

$$\sum_{m=1}^{M_f} \omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2)$$

(ここで、

【数2】

$$f(x; \mu_{f(m)}, \sigma_{f(m)}^2) \quad (m = 1, 2, \dots, M_f)$$

はガウス分布を表し、

【数3】

 x

は入力データを表す) における統計量である混合重み係数

【数4】

$$\omega_{f(m)} \quad (m = 1, 2, \dots, M_f)$$

、平均値

【数5】

$$\mu_{f(m)} \quad (m = 1, 2, \dots, M_f)$$

および分散値

【数6】

$$\sigma_{f(m)}^2 \quad (m = 1, 2, \dots, M_f)$$

の少なくとも1つを、前記 N_g ($N_g \geq 1$) の参照モデル

【数7】

$$\sum_{l=1}^{L_{g(i)}} v_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \quad (i = 1, 2, \dots, N_g)$$

(ここで、

【数 8】

$$g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \quad (i = 1, 2, \dots, N_g, l = 1, 2, \dots, L_{(i)})$$

はガウス分布を表し、

【数 9】

$$L_{g(i)} \quad (i = 1, 2, \dots, N_g)$$

は各参照モデルの混合分布数を表し、

【数 10】

$$\nu_{g(i,l)} \quad (l = 1, 2, \dots, L_{g(i)})$$

は混合重み係数を表し、

【数 11】

$$\mu_{g(i,l)} \quad (l = 1, 2, \dots, L_{g(i)})$$

は平均値を表し、

【数 12】

$$\sigma_{g(i,l)}^2 \quad (l = 1, 2, \dots, L_{g(i)})$$

は分散値を表す) に対する前記標準モデルの尤度

【数 13】

$$\log P = \sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \log \left[\sum_{m=1}^{M_f} \omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2) \right] \left\{ \sum_{l=1}^{L_{g(i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx$$

を最大化または極大化するように、計算する

ことを特徴とする請求項 24 記載の標準モデル作成装置。

【請求項 31】 前記統計量推定部は、混合分布数が M_f 個 ($M_f \geq 1$) の混合ガウス分布により表現された前記標準モデル

【数 1 4】

$$\sum_{m=1}^{M_f} \omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2)$$

(ここで、

【数 1 5】

$$f(x; \mu_{f(m)}, \sigma_{f(m)}^2) \quad (m = 1, 2, \dots, M_f)$$

はガウス分布を表し、

【数 1 6】

$$x = (x_{(1)}, x_{(2)}, \dots, x_{(J)}) \in R^J$$

は J (≥ 1) 次元の入力データを表す) における統計量である混合重み係数

【数 1 7】

$$\omega_{f(m)} \quad (m = 1, 2, \dots, M_f)$$

、J (≥ 1) 次元の平均値

【数 1 8】

$$\mu_{f(m)} = (\mu_{f(m,1)}, \mu_{f(m,2)}, \dots, \mu_{f(m,J)}) \in R^J \quad (m = 1, 2, \dots, M_f)$$

および J (≥ 1) 次元の分散値 (共分散行列の J 個の対角成分)

【数 1 9】

$$\sigma_{f(m)}^2 = (\sigma_{f(m,1)}^2, \sigma_{f(m,2)}^2, \dots, \sigma_{f(m,J)}^2) \in R^J \quad (m = 1, 2, \dots, M_f)$$

の少なくとも 1 つを、

【数 2 0】

$$\omega_{f(m)} = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{L_g(i)} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{k=1}^{M_f} \sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, k) \left\{ \sum_{l=1}^{L_g(i)} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f)$$

【数 2 1】

$$\mu_{f(m,j)} = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) x_{(j)} \left\{ \sum_{l=1}^{L_g(i)} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{L_g(i)} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【数 2 2】

$$\sigma_{f(m,j)}^2 = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) (x_{(j)} - \mu_{f(m,j)})^2 \left\{ \sum_{l=1}^{L_g(i)} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{L_g(i)} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

(ここで、

【数 2 3】

$$\gamma(x, m) = \frac{\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2)}{\sum_{k=1}^{M_f} \omega_{f(k)} f(x, \mu_{f(k)}, \sigma_{f(k)}^2)} \quad (m = 1, 2, \dots, M_f)$$

である)を利用して1以上繰り返して計算すること
ことを特徴とする請求項30記載の標準モデル作成装置。

【請求項32】 前記統計量推定部は、

【数24】

$$\gamma(x, m) = \frac{\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2)}{\sum_{k=1}^{M_f} \omega_{f(k)} f(x, \mu_{f(k)}, \sigma_{f(k)}^2)} \quad (m = 1, 2, \dots, M_f)$$

における出力分布

【数25】

$$\omega_{f(k)} f(x, \mu_{f(k)}, \sigma_{f(k)}^2) \quad (k = 1, 2, \dots, M_f)$$

の中から、前記数24における出力分布

【数26】

$$\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2) \quad (m = 1, 2, \dots, M_f)$$

との分布間距離が近い $P_h(m)$ 個 ($1 \leq P_h(m) \leq M_f$) の出力分布

【数27】

$$\omega_{f(m,p)} f(x, \mu_{f(m,p)}, \sigma_{f(m,p)}^2) \quad (m = 1, 2, \dots, M_f, p = 1, 2, \dots, P_{h(m)})$$

(ここで、

【数28】

$$\mu_{f(m,p)} = (\mu_{f(m,p,1)}, \mu_{f(m,p,2)}, \dots, \mu_{f(m,p,J)}) \in R^J$$

は、 $J (\geq 1)$ 次元の平均値であり、

【数29】

$$\sigma_{f(m,p)}^2 = (\sigma_{f(m,p,1)}^2, \sigma_{f(m,p,2)}^2, \dots, \sigma_{f(m,p,J)}^2) \in R^J$$

は、 $J (\geq 1)$ 次元の分散値 (共分散行列の J 個の対角成分) である) を選択

し、選択した $P_{h(m)}$ ($m=1, 2, \dots, M_f$) 個の出力分布を用いて、単一の重み付きガウス分布

【数 3 0】

$$u_{h(m)} h(x; \mu_{h(m)}, \sigma_{h(m)}^2) \quad (m=1, 2, \dots, M_f)$$

(ここで、

【数 3 1】

$$\mu_{h(m)} = (\mu_{h(m,1)}, \mu_{h(m,2)}, \dots, \mu_{h(m,J)}) \in R^J$$

は、 J (≥ 1) 次元の平均値であり、

【数 3 2】

$$\sigma_{h(m)}^2 = (\sigma_{h(m,1)}^2, \sigma_{h(m,2)}^2, \dots, \sigma_{h(m,J)}^2) \in R^J$$

は、 J (≥ 1) 次元の分散値 (共分散行列の J 個の対角成分) である) における重み、平均値および分散値 (共分散行列の対角成分) を、それぞれ、

【数 3 3】

$$u_{h(m)} = \sum_{p=1}^{P_{h(m)}} \omega_{f(m,p)} \quad (m=1, 2, \dots, M_f)$$

【数 3 4】

$$\mu_{h(m,j)} = \frac{\sum_{p=1}^{P_{h(m)}} \omega_{f(m,p)} \mu_{f(m,p,j)}}{\sum_{p=1}^{P_{h(m)}} \omega_{f(m,p)}} \quad (m=1, 2, \dots, M_f, j=1, 2, \dots, J)$$

【数 3 5】

$$\sigma_{h(m,j)}^2 = \frac{\sum_{p=1}^{P_{f(m)}} \omega_{f(m,p)} (\sigma_{f(m,p,j)}^2 + \mu_{f(m,p,j)}^2)}{\sum_{p=1}^{P_{h(m)}} \omega_{f(m,p)}} - \mu_{h(m,j)}^2$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

を利用して算出して、

前記数 2 7 における $P_h(m)$ 個の出力分布の近傍の

【数 3 6】

x

に対して、

【数 3 7】

$$\gamma(x, m) = \frac{\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2)}{\sum_{k=1}^{M_f} \omega_{f(k)} f(x, \mu_{f(k)}, \sigma_{f(k)}^2)} \approx \frac{\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2)}{u_{h(m)} h(x, \mu_{h(m)}, \sigma_{h(m)}^2)}$$

$$(m = 1, 2, \dots, M_f)$$

による近似式を利用して、上記範囲以外の

【数 3 8】

x

に対して、

【数 3 9】

$$\gamma(x, m) = \frac{\omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2)}{\sum_{k=1}^{M_f} \omega_{f(k)} f(x; \mu_{f(k)}, \sigma_{f(k)}^2)} \approx 0.0$$

$$(m = 1, 2, \dots, M_f)$$

による近似式を利用して前記統計量を計算する

ことを特徴とする請求項 3 1 記載の標準モデル作成装置。

【請求項 3 3】 前記出力分布

【数 4 0】

$$\omega_{f(m,p)} f(x; \mu_{f(m,p)}, \sigma_{f(m,p)}^2) \quad (m = 1, 2, \dots, M_f, p = 1, 2, \dots, P_{h(m)})$$

の近傍の

【数 4 1】

x

とは、前記数 4 0 が示す出力分布との分布間距離が近い $Q_g(m, i)$ 個 ($1 \leq Q_g(m, i) \leq L_g(i)$) の前記参照ベクトルの出力分布

【数 4 2】

$$v_{g(i,l)} g(x; \mu_{g(l)}, \sigma_{g(l)}^2) \quad (i = 1, 2, \dots, N_g, l = 1, 2, \dots, Q_{g(m,i)})$$

が存在する空間であって、前記統計量推定部は、混合分布数が M_f 個 ($M_f \geq 1$) の混合ガウス分布により表現された前記標準モデル

【数 4 3】

$$\sum_{m=1}^{M_f} \omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2)$$

(ここで、

【数 4 4】

$$f(x; \mu_{f(m)}, \sigma_{f(m)}^2) \quad (m = 1, 2, \dots, M_f)$$

はガウス分布を表し、

【数 4 5】

$$x = (x_{(1)}, x_{(2)}, \dots, x_{(J)}) \in R^J$$

は J (≥ 1) 次元の入力データを表す) における統計量である混合重み係数

【数 4 6】

$$\omega_{f(m)} \quad (m = 1, 2, \dots, M_f)$$

、 J (≥ 1) 次元の平均値

【数 4 7】

$$\mu_{f(m)} = (\mu_{f(m,1)}, \mu_{f(m,2)}, \dots, \mu_{f(m,J)}) \in R^J \quad (m = 1, 2, \dots, M_f)$$

および J (≥ 1) 次元の分散値 (共分散行列の J 個の対角成分)

【数 4 8】

$$\sigma_{f(m)}^2 = (\sigma_{f(m,1)}^2, \sigma_{f(m,2)}^2, \dots, \sigma_{f(m,J)}^2) \in R^J \quad (m = 1, 2, \dots, M_f)$$

の少なくとも 1 つを、

【数 4 9】

$$\omega_{f(m)} = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{Q_{g(m,l)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{k=1}^{M_f} \sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, k) \left\{ \sum_{l=1}^{Q_{g(m,l)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f).$$

【数 5 0】

$$\mu_{f(m,j)} = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) x_{(j)} \left\{ \sum_{l=1}^{Q_{g(m,i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{Q_{g(m,i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【数 5 1】

$$\sigma_{f(m,j)}^2 = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) (x_{(j)} - \mu_{f(m,j)})^2 \left\{ \sum_{l=1}^{Q_{g(m,i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{Q_{g(m,i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

を利用して R 回 ($R \geq 1$) 繰り返して計算する

ことを特徴とする請求項 3 2 記載の標準モデル作成装置。

【請求項 3 4】 前記出力分布

【数 5 2】

$$\omega_{f(m,p)} f(x; \mu_{f(m,p)}, \sigma_{f(m,p)}^2) \quad (m = 1, 2, \dots, M_f, p = 1, 2, \dots, P_{h(m)})$$

との分布間距離が近い $Q_g(m, i)$ 個 ($1 \leq Q_g(m, i) \leq L_g(i)$)

の前記参照ベクトルの出力分布とは、前記参照モデルの出力分布

【数 5 3】

$$\nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \quad (i = 1, 2, \dots, N_g, l = 1, 2, \dots, L_{g(i)})$$

のうち分布間距離が第 1 番目から第 G ($1 \leq G \leq M_f$) 番目 (G: 近傍指示パラメータ) までに近い前記標準モデルの出力分布が前記数 5 2 である前記参照ベ

クトルの出力分布である

ことを特徴とする請求項 33 記載の標準モデル作成装置。

【請求項 35】 前記統計量推定部は、前記近似を繰り返し、その繰り返し回数に対応して前記 G を変化させる

ことを特徴とする請求項 34 に記載の標準モデル作成装置。

【請求項 36】 前記統計量推定部は、前記繰り返し回数に対応して前記 G を減少させる

ことを特徴とする請求項 35 記載の標準モデル作成装置。

【請求項 37】 前記統計量推定部は、前記単一のガウス分布として、前記標準モデルの出力分布

【数 54】

$$\omega_{f(k)} f(x, \mu_{f(k)}, \sigma_{f(k)}^2)$$

に分布間距離の最も近い 1 個の出力分布 ($P_h(m) = 1, m = 1, 2, \dots, M_f$) である前記数 54 を用いて、

【数 55】

$$\gamma(x, m) \approx \frac{\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2)}{u_{h(m)} h(x, \mu_{h(m)}, \sigma_{h(m)}^2)} \approx 1.0$$

$$(m = 1, 2, \dots, M_f)$$

を利用して第 2 近似を行う

ことを特徴とする請求項 31～36 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 38】 前記統計量推定部は、前記近似を繰り返し、その繰り返し回数に対応して前記 $P_h(m)$ を変化させる

ことを特徴とする請求項 32 に記載の標準モデル作成装置。

【請求項 39】 前記統計量推定部は、前記繰り返し回数に対応して前記 $P_h(m)$ を減少させる

ことを特徴とする請求項 38 記載の標準モデル作成装置。

【請求項 40】 前記事象は、隠れマルコフモデルの状態である

ことを特徴とする請求項 1～39 のいずれか 1 項に記載の標準モデル作成装置

。

【請求項 41】 事象の出力確率によって定義される認識用のモデルである標準モデルを作成する方法であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成ステップと、

外部から参照モデルを取得して参照モデル記憶手段に格納すること、および、参照モデルを作成して前記参照モデル記憶手段に格納することの少なくとも一方を行う参照モデル準備ステップと、

認識の対象に関する情報である利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する参照モデル選択ステップとを含み、

前記標準モデル作成ステップでは、前記参照モデル選択ステップで選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする標準モデル作成方法。

【請求項 42】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識用のモデルである標準モデルを作成する方法であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成ステップと、

外部から参照モデルを取得して参照モデル記憶手段に格納すること、および、参照モデルを作成して前記参照モデル記憶手段に格納することの少なくとも一方を行う参照モデル準備ステップと、

認識の対象に関する情報である利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する参照モデル選択ステップとを含み、

前記標準モデル作成ステップでは、前記参照モデル選択ステップで選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする標準モデル作成方法。

【請求項 43】 事象の出力確率によって定義される認識用のモデルである標準モデルを作成する方法であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成ステップと、

認識の対象に関する情報である利用情報に基づいて、参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する参照モデル選択ステップとを含み、

前記標準モデル作成ステップでは、前記参照モデル選択ステップで選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

を含むことを特徴とする標準モデル作成方法。

【請求項 44】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識用のモデルである標準モデルを作成する方法であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成ステップ

認識の対象に関する情報である利用情報に基づいて、参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する参照モデル選択ステップとを含み、

前記標準モデル作成ステップでは、前記参照モデル選択ステップで選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

を含むことを特徴とする標準モデル作成方法。

【請求項 4 5】 前記標準モデル作成方法は、さらに、

認識の対象に関する情報である利用情報を作成する利用情報作成ステップを含み、

前記参照モデル選択ステップでは、作成された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する

ことを特徴とする請求項 4 1 ～ 4 4 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 4 6】 前記標準モデル作成方法は、さらに、

認識の対象に関する情報である利用情報を端末装置から受信する利用情報受信ステップを含み、

前記参照モデル選択ステップでは、作成された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から 1 以上の参照モデルを選択する

ことを特徴とする請求項 4 1 ～ 4 4 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 4 7】 前記標準モデル作成ステップでは、作成する標準モデルの仕様に関する情報である仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 4 1 ～ 4 6 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 4 8】 事象の出力確率によって定義される認識用のモデルである標準モデルを作成する方法であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデル

に対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成ステップを含み、

前記標準モデル作成ステップでは、作成する標準モデルの仕様に関する情報である仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

を含むことを特徴とする標準モデル作成方法。

【請求項 49】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識用のモデルである標準モデルを作成する方法であって、

特定の対象を認識するために予め作成されたモデルである 1 以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成ステップを含み、

前記標準モデル作成ステップでは、作成する標準モデルの仕様に関する情報である仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

を含むことを特徴とする標準モデル作成方法。

【請求項 50】 前記仕様情報は、標準モデルを使用するアプリケーションプログラムの種類に対応づけられた仕様を示す

ことを特徴とする請求項 47～49 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 51】 前記標準モデル作成ステップでは、標準モデルを使用するアプリケーションプログラムと標準モデルの仕様との対応を示すアプリケーション仕様対応データベースから、起動されるアプリケーションプログラムに対応する仕様を読み出し、読み出した仕様に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 47～49 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 5 2】 作成する標準モデルの仕様に関する情報である仕様情報を作成する仕様情報作成ステップを含み、

前記標準モデル作成ステップでは、前記仕様情報作成ステップで作成した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 4 7～4 9 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 5 3】 前記標準モデル作成方法は、さらに、

作成する標準モデルの仕様に関する情報である仕様情報を端末装置から受信する仕様情報受信ステップを含み、

前記標準モデル作成ステップでは、前記仕様情報受信ステップで受信した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 4 7～4 9 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 5 4】 前記参照モデルおよび前記標準モデルは、1 以上のガウス分布を用いて表現され、

前記標準モデル作成ステップでは、前記仕様情報に基づいて、前記標準モデルの構造として、少なくとも前記標準モデルの混合分布数を決定する

ことを特徴とする請求項 4 7～5 3 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 5 5】 前記参照モデルおよび前記標準モデルは、1 以上のガウス分布を用いて表現され、

前記参照モデル記憶手段には、少なくとも 1 対の参照モデルの混合分布数が異なる参照モデルが記憶され、

前記標準モデル作成ステップでは、少なくとも 1 対の参照モデルの混合分布数が異なる参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 4 1～5 4 のいずれか 1 項に記載の標準モデル作成方

法。

【請求項 56】 前記標準モデル作成方法は、さらに、

前記標準モデル作成ステップで作成した標準モデルを端末装置に送信する標準モデル送信ステップを含む

ことを特徴とする請求項 41～55 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 57】 前記標準モデル作成方法は、さらに、

端末装置から送信される参照モデルを受信する参照モデル受信ステップを含み

前記標準モデル作成ステップでは、少なくとも前記参照モデル受信ステップで受信した参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算する

ことを特徴とする請求項 41～56 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 58】 前記参照モデル準備ステップでは、さらに、前記参照モデル記憶手段で記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 41 または 42 記載の標準モデル作成方法。

【請求項 59】 前記参照モデル準備ステップでは、認識の対象に関する情報である利用情報、および作成する標準モデルの仕様に関する情報である仕様情報の少なくとも一方に基づいて、前記参照モデル記憶手段で記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 58 記載の標準モデル作成方法。

【請求項 60】 前記標準モデル作成方法は、さらに、認識の対象に関する情報である利用情報、および作成する標準モデルの仕様に関する情報である仕様情報の少なくとも一方と、前記参照モデル記憶手段に記憶された参照モデルに基づいて、前記利用情報および前記仕様情報の少なくとも一方と前記参照モデルとの類似度情報を作成する類似度情報作成ステップを含み、

前記参照モデル準備ステップでは、前記類似度情報作成ステップで作成した類似度情報に基づいて、前記参照モデル記憶手段で記憶する参照モデルの更新および追加の少なくとも一方を行う

び追加の少なくとも一方を行うか否かを決定する

ことを特徴とする請求項 58 または 59 記載の標準モデル作成方法。

【請求項 61】 前記標準モデル作成方法は、さらに、
端末装置から送信される参照モデルを受信する参照モデル受信ステップを含み

、
前記参照モデル準備ステップでは、前記参照モデル受信ステップで受信した参照モデルを用いて前記参照モデル記憶手段で記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 58 ～ 60 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 62】 前記標準モデル作成ステップでは、
作成する標準モデルの構造を決定する標準モデル構造決定ステップと、
構造が決定された前記標準モデルを特定する統計量の初期値を決定する初期標準モデル作成ステップと、

前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を推定して計算する統計量推定ステップとを有する

ことを特徴とする請求項 41 ～ 61 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 63】 前記初期標準モデル作成ステップでは、前記統計量推定ステップが標準モデルの統計量を計算するために用いる、1 以上の前記参照モデルを用いて前記標準モデルを特定する統計量の初期値を決定する

ことを特徴とする請求項 62 記載の標準モデル作成方法。

【請求項 64】 前記初期標準モデル作成ステップでは、標準モデルの種類を識別するクラス ID に基づいて、前記初期値を決定する

ことを特徴とする請求項 62 記載の標準モデル作成方法。

【請求項 65】 前記初期標準モデル作成ステップでは、前記参照モデルからクラス ID を特定し、特定したクラス ID に対応づけられた初期値を前記初期値と決定する

ことを特徴とする請求項 6 4 記載の標準モデル作成方法。

【請求項 6 6】 前記初期標準モデル作成ステップでは、前記クラス ID と前記初期値と前記参照モデルとの対応を示す対応表に従って、前記初期値を決定する

ことを特徴とする請求項 6 5 記載の標準モデル作成方法。

【請求項 6 7】 前記初期標準モデル作成ステップでは、前記クラス ID が対応づけられた初期値であるクラス ID 付き初期標準モデル、または、前記クラス ID が対応づけられた参照モデルであるクラス ID 付き参照モデルを作成または外部から取得することによって、前記対応表を生成する

ことを特徴とする請求項 6 6 記載の標準モデル作成方法。

【請求項 6 8】 前記事象は、隠れマルコフモデルの状態である

ことを特徴とする請求項 4 1 ～ 6 7 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 6 9】 標準モデルを作成する方法であって、

請求項 4 1 ～ 6 8 のいずれか 1 項に記載の標準モデル作成方法に含まれるステップをコンピュータに実行させる

ことを特徴とするプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、隠れマルコフモデル、ベイジ理論、線形判別分析などの確率モデルによる音声認識、文字認識、画像認識などのパターン認識、ベイジアンネットなどの確率モデルによる意図理解（意図の認識）、確率モデルによるデータマイニング（データ特性の認識）、確率モデルによる人物検出、指紋認証、顔認証、虹彩認証（対象を認識して特定の対象かどうかを判断する）、株価予測、天気予測などの予測（状況を認識して判断する）などに用いられる標準モデルの作成装置およびその方法に関する。

【0002】

近年、インターネットなどの普及により、ネットワークの大容量化、通信コス

トの低価格化が進んでいる。このため、ネットワークを利用することで、多くの認識用モデル（参照モデル）を収集することが可能となってきた。例えば、音声認識において、様々な研究機関で配布している多くの音声認識用モデル（子供用モデル、成人用モデル、高齢者用モデル、自動車内用モデル、携帯電話用モデルなど）をインターネットによりダウンロードすることが可能となってきた。また、ネットワークによる機器連携により、カーナビゲーションシステムなどで利用する音声認識用モデルをテレビやパソコンなどにダウンロードできるようになってきている。また、意図理解において、各地の様々な人の経験を学習した認識用モデルを、ネットワークを通して収集することが可能となってきた。

【0003】

また、認識技術の発展により、認識用モデルは、パソコン、テレビのリモコン、携帯電話、カーナビゲーションシステムなど、CPUパワー、メモリ量などの仕様の異なる幅広い機器に利用されるようになってきている。また、セキュリティなどの認識精度が要求されるアプリケーションや、テレビのリモコンでの操作のように認識結果が出力されるまでの時間の速さが要求されるアプリケーションなど、要求仕様の異なる幅広いアプリケーションに利用されるようになってきている。

【0004】

また、認識技術は、認識対象の異なる多くの環境で利用されるようになってきている。例えば、音声認識において、子供の声、成人の声、高齢者の声を認識したり、自動車内での声、携帯電話での声を認識するなど、多くの環境で利用される。

【0005】

これらの社会環境の変化を鑑みると、多くの認識用モデル（参照モデル）を有効に活用することで、機器やアプリケーションの仕様、利用環境に適した精度の高い認識用モデル（標準モデル）を短時間に作成して利用者に提供することが望まれると考えられる。

【0006】

【従来の技術】

音声認識などのパターン認識の分野で、認識用の標準モデルとして確率モデルを用いる方法が近年注目されており、特に、隠れマルコフモデル（以下HMMと呼ぶ）や混合ガウス分布モデル（以下GMMと呼ぶ）が広く用いられている。また、意図理解において、意図、知識、嗜好などを表す標準モデルとして確率モデルを用いる方法が近年注目されており、特に、ベイジアンネットなどが広く用いられている。また、データマイニングの分野で、データを分類するために各カテゴリの代表モデルとして確率モデルを用いる方法が注目されており、GMMなどが広く用いられている。また、音声認証、指紋認証、顔認証、虹彩認証などの認証の分野で、認証用の標準モデルとして確率モデルを用いる方法が注目されており、GMMなどが用いられている。HMMにより表現される標準モデルの学習アルゴリズムとしてバウム・ウェルチ（B a u m - W e l c h）の再推定の方法が広く用いられている（例えば、今井聖著、“音声認識”、p p . 1 5 0 - 1 5 2、共立出版株式会社、1995年11月25日発行参照）。また、GMMにより表現される標準モデルの学習アルゴリズムとしてEM（E x p e c t a t i o n - M a x i m i z a t i o n）アルゴリズムが広く用いられている（例えば、古井貞▲ひろ▼著、“音声情報処理”、p p . 1 0 0 - 1 0 4、森北出版株式会社、1998年6月30日発行参照）。EMアルゴリズムでは、標準モデル

【0007】

【数56】

$$\sum_{m=1}^{M_f} \omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2)$$

【0008】

(ここで、

【0009】

【数57】

$$f(x; \mu_{f(m)}, \sigma_{f(m)}^2) \quad (m = 1, 2, \dots, M_f)$$

【0010】

はガウス分布を表し、

【0011】

【数58】

$$x = (x_{(1)}, x_{(2)}, \dots, x_{(J)}) \in R^J$$

【0012】

はJ (≥1) 次元の入力データを表す) における統計量である混合重み係数

【0013】

【数59】

$$\omega_{f(m)} \quad (m = 1, 2, \dots, M_f)$$

【0014】

、J (≥1) 次元の平均値

【0015】

【数60】

$$\mu_{f(m)} = (\mu_{f(m,1)}, \mu_{f(m,2)}, \dots, \mu_{f(m,J)}) \in R^J$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0016】

およびJ (≥1) 次元の分散値 (共分散行列のJ個の対角成分)

【0017】

【数61】

$$\sigma_{f(m)}^2 = (\sigma_{f(m,1)}^2, \sigma_{f(m,2)}^2, \dots, \sigma_{f(m,J)}^2) \in R^J$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0018】

を、

N個の学習データ

【0019】

【数62】

$$x[i] = (x_{(1)}[i], x_{(2)}[i], \dots, x_{(J)}[i]) \in R^J \quad (i = 1, 2, \dots, N)$$

【0020】

を用いて、学習データに対する尤度

【0021】

【数63】

$$\log P = \sum_{i=1}^N \log \left[\sum_{m=1}^{M_f} \omega_{f(m)} f(x[i]; \mu_{f(m)}, \sigma_{f(m)}^2) \right]$$

【0022】

を最大化もしくは極大化するように、

【0023】

【数64】

$$\omega_{f(m)} = \frac{\sum_{i=1}^N \gamma(x[i], m)}{\sum_{k=1}^{M_f} \sum_{i=1}^N \gamma(x[i], k)}$$

$$(m = 1, 2, \dots, M_f)$$

【0024】

【数 6 5】

$$\mu_{f(m,j)} = \frac{\sum_{i=1}^N \gamma(x[i], m) x_{(j)}}{\sum_{i=1}^N \gamma(x[i], m)}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0025】

【数 6 6】

$$\sigma_{f(m,j)}^2 = \frac{\sum_{i=1}^N \gamma(x[i], m) (x_{(j)} - \mu_{f(m,j)})^2}{\sum_{i=1}^N \gamma(x[i], m)}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0026】

(ここで、

【0027】

【数 6 7】

$$\gamma(x[i], m) = \frac{\omega_{f(m)} f(x[i]; \mu_{f(m)}, \sigma_{f(m)}^2)}{\sum_{k=1}^{M_f} \omega_{f(k)} f(x[i]; \mu_{f(k)}, \sigma_{f(k)}^2)} \quad (m = 1, 2, \dots, M_f)$$

【0028】

である)を利用して1以上繰り返して計算して学習を行う。また、ベイズ推定法(例えば、繁樹算男著、"ベイズ統計入門"、pp. 42-53、東京大学出版会、1985年4月30日発行参照)などの方法も提案されている。バウム・ウェルチの再推定の方法、EMアルゴリズム、ベイズ推定法のいずれの学習アルゴリズムも、学習データに対する確率(尤度)を最大化もしくは極大化するように

標準モデルのパラメータ（統計量）を計算して標準モデルを作成する。これらの学習方法では、確率（尤度）を最大化もしくは極大化するという数学的な最適化が実現されている。

【0029】

上記の学習方法を音声認識の標準モデルの作成に用いた場合、多様な話者や雑音などの音響的特徴量の変動に対応するために多数の音声データで標準モデルを学習することが望ましい。また、意図理解に用いた場合、多様な話者や状況などの変動に対応するために多数のデータで標準モデルを学習することが望ましい。また、虹彩認証に用いた場合、太陽光、カメラ位置・回転などの変動に対応するために多数の虹彩画像データで標準モデルを学習することが望ましい。しかしながら、このような多量のデータを取り扱う場合、学習に膨大な時間がかかるため、利用者に標準モデルを短時間に提供できない。また、多量のデータを蓄積するためのコストが膨大となる。また、ネットワークを利用してデータを収集した場合、通信コストが膨大となる。

【0030】

一方、複数のモデル（以下、標準モデルの作成のために参照用として準備されるモデルを「参照モデル」と呼ぶ。）を合成することで標準モデルを作成する方法が提案されている。参照モデルは、多くの学習データを確率分布の母数（平均、分散など）で表現した確率分布モデルであり、多くの学習データの特徴を少数のパラメータ（母数）で集約したものであるため、参照モデルを用いて標準モデルを作成した場合、標準モデルを短時間に作成することが可能となる。ここでのモデルはガウス分布で表現されている。

【0031】

第1の従来方法では、参照モデルはGMMで表現されており、複数の参照モデルのGMMを重み付きで合成することで標準モデルを作成している（例えば、特許文献1参照）。

【0032】

また、第2の従来方法では、第1の従来方式に加えて、学習データに対する確率（尤度）を最大化あるいは極大化して線形結合された混合重みを学習すること

で標準モデルを作成している（例えば、特許文献2参照）。

【0033】

また、第3の従来方法では、標準モデルの平均値を参照モデルの平均値の線形結合で表現し、入力データに対する確率（尤度）を最大化あるいは極大化して線形結合係数を学習することで標準モデルを作成している。ここでは学習データとして特定話者の音声データを用いており標準モデルを音声認識用の話者適応モデルとして用いている（例えば、非特許文献1参照）。

【0034】

また、第4の従来方法では、参照モデルは単一ガウス分布で表現されており、複数の参照モデルのガウス分布を合成したのちに、クラスタリングにより同一クラスに属するガウス分布を統合することで標準モデルを作成している（例えば、特許文献3参照）。

【0035】

また、第5の従来方法では、複数の参照モデルは同数の混合数の混合ガウス分布で表現され、各ガウス分布には通し番号が付与されている。標準モデルは、同一の通し番号をもつガウス分布を合成することにより作成される。合成する複数の参照モデルは利用者に音響的に近い話者で作成されたモデルであり、作成させる標準モデルは話者適応モデルである（例えば、非特許文献2参照）。

【0036】

【特許文献1】

特開平4-125599号公報（第3項、第1図）

【0037】

【特許文献2】

特開平10-268893号公報（第3-6頁、第1図）

【0038】

【特許文献3】

特開平9-81178号公報（第3-4項、第4図、第5図）

【0039】

【非特許文献1】

M. J. F. Gales, "Cluster Adaptive Training For Speech Recognition", 1998年、ICSLP98予稿集、pp. 1783-1786

【0040】

【非特許文献2】

芳澤、外6名、"十分統計量と話者距離を用いた音韻モデルの教師なし学習法"、2002年3月1日、電子情報通信学会、Vol. J85-D-II、No. 3、pp. 382-389

【0041】

【発明が解決しようとする課題】

しかしながら、第1の従来方法では、合成する参照モデル数の増加とともに標準モデルの混合数が増加して、標準モデルのための記憶容量、認識処理量が膨大となり実用的でない。また、仕様に応じて標準モデルの混合数を制御することができない。また、標準モデルの作成において、最尤推定などの最適化アルゴリズムによる学習ではなく、標準モデルは参照モデルの単純な混合和であるため、高精度の標準モデルが作成できない。これらの課題は、合成する参照モデルの数の増加に伴い顕著になってくると考えられる。

【0042】

第2の従来方法では、合成する参照モデル数の増加とともに標準モデルの混合数が増加して、標準モデルのための記憶容量、認識処理量が膨大となり実用的でない。また、仕様に応じて標準モデルの混合数を制御することができない。また、標準モデルは、参照モデルの単純な混合和であり学習するパラメータが混合重みに限定されているため、高精度の標準モデルが作成できない。また、標準モデルの作成において、学習データを用いて学習を行っているため学習時間がかかる。これらの課題は、合成する参照モデルの数の増加に伴い顕著になってくると考えられる。

【0043】

第3の従来方法では、学習するパラメータが参照モデルの平均値の線形結合係数に限定されているため高精度の標準モデルが作成できない。また、標準モデル

の作成において、学習データを用いて学習を行っているため学習時間がかかる。

【0044】

第4の従来方法では、クラスタリングをヒューリスティックに行うため高精度の標準モデルを作成することが困難である。また、参照モデルは単一のガウス分布であるため精度が低く、それらを統合した標準モデルの精度は低い。認識精度に関する課題は、合成する参照モデルの数の増加に伴い顕著になってくると考えられる。

【0045】

第5の従来方法では、標準モデルは、同一の通し番号をもつガウス分布を合成することにより作成されるが、最適な標準モデルを作成するためには、一般的には合成するガウス分布は1対1に対応するとは限らないため、認識精度が低下する。また、複数の参照モデルが異なる混合数をもつ場合に標準モデルを作成することができない。また、一般的には、参照モデルにおけるガウス分布に通し番号が付与されておらず、この場合に標準モデルを作成することができない。また、仕様に応じて標準モデルの混合数を制御することができない。

【0046】

そこで、本発明は、このような問題点に鑑みてなされたものであり、隠れマルコフモデル、ベイズ理論、線形判別分析などの確率モデルによる音声認識、文字認識、画像認識などのパターン認識、ベイジアンネットなどの確率モデルによる意図理解（意図の認識）、確率モデルによるデータマイニング（データ特性の認識）、株価予測、天気予測などの予測（状況を認識して判断する）などに用いられる高精度な標準モデルを作成する標準モデル作成装置等を提供することを目的とする。

【0047】

また、本発明は、学習のためのデータや教師データを必要とすることなく、簡易に標準モデルを作成することが可能な標準モデル作成装置等を提供することをも目的とする。

【0048】

さらに、本発明は、標準モデルを利用する認識の対象にふさわしい標準モデル

を作成したり、標準モデルを用いて認識処理を実行する装置の仕様や環境に適した標準モデルを作成することが可能な汎用性および柔軟性に優れた標準モデル作成装置等を提供することを目的とする。

【0049】

本発明で用いる「認識」とは、音声認識などの狭義の意味での認識だけではなく、パターンマッチング、識別、認証、ベイズ推定や予測など、確率で表現された標準モデルを利用するものの全般を意味する。

【0050】

【課題を解決するための手段】

上記目的を達成するために、本発明に係る標準モデル作成装置は、事象の集合の出力確率によって定義される認識用のモデルである標準モデル、あるいは、事象の集合と事象または事象間の遷移の出力確率とによって定義される認識用のモデルである標準モデルを作成する装置であって、特定の対象を認識するために予め作成されたモデルである参照モデルを1以上記憶する参照モデル記憶手段と、前記参照モデル記憶手段に記憶された1以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段とを備えることを特徴とする。

【0051】

これによって、1以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように標準モデルの統計量が計算され、標準モデルが作成されるので、音声データ等の学習データや教師データを必要とすることなく簡易に標準モデルが作成されるとともに、既に作成された複数の参照モデルを総合的に勘案した高精度な標準モデルが作成される。

【0052】

ここで、前記標準モデル作成装置は、さらに、外部から参照モデルを取得して前記参照モデル記憶手段に格納すること、および、参照モデルを作成して前記参照モデル記憶手段に格納することの少なくとも一方を行う参照モデル準備手段を備えてもよい。

【0053】

これによって、標準モデル作成装置の外部から新たな参照モデルを取り込み、取り込んだ参照モデルに基づいた標準モデルの作成が可能となるので、様々な認識対象に対応した汎用性の高い標準モデル作成装置が実現される。

【0054】

また、前記標準モデル作成装置は、さらに、認識の対象に関する情報である利用情報を作成する利用情報作成手段と、作成された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から1以上の参照モデルを選択する参照モデル選択手段とを備え、前記標準モデル作成手段は、前記参照モデル選択手段が選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

【0055】

これによって、利用者の特徴、利用者の年齢、性別、利用環境などの利用情報に基づいて、準備された複数の参照モデルの中から認識対象に適した参照モデルだけが選択され、それら参照モデルを統合した標準モデルが作成されるので、認識対象により特化した精度の高い標準モデルが作成される。

【0056】

ここで、前記標準モデル作成装置は、さらに、前記利用情報と選択された参照モデルに関する情報との類似度を算出して、前記類似度が所定のしきい値以上であるか否かを判定して判定信号を作成する類似度判定手段を備えてもよい。これによって、利用情報にふさわしい（近い）参照モデルが参照モデル記憶手段に存在しない場合に、参照モデルの準備の要求を行うことができる。

【0057】

また、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、認識の対象に関する情報である利用情報を前記端末装置から受信する利用情報受信手段と、受信された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から1以上の参照モデルを選択する参照モデル選択手段とを備え、前記標準モデル作成手段は、前記参照モデル選択手段が選択した参照モデルに対する前記標準モデルの確率また

は尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

【0058】

これによって、通信路を介して送信されてきた利用情報に基づいて標準モデルが作成されるので、遠隔制御による標準モデルの生成が可能になるとともに、通信システムを基盤とする認識システムの構築が実現される。

【0059】

また、前記標準モデル作成装置は、さらに、作成する標準モデルの仕様に関する情報である仕様情報を作成する仕様情報作成手段を備え、前記標準モデル作成手段は、前記仕様情報作成手段が作成した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

【0060】

これによって、標準モデルを使用する装置のCPUパワー、記憶容量、要求される認識精度、要求される認識処理時間などの仕様情報に基づいて標準モデルが作成されるので、特定の仕様条件を満たす標準モデルの生成が可能となり、計算エンジン等の認識処理に必要なリソース環境に適した標準モデルの生成が実現される。

【0061】

ここで、前記仕様情報は、例えば、標準モデルを使用するアプリケーションプログラムの種類に対応づけられた仕様を示すような情報であってもよい。そして、前記標準モデル作成装置は、さらに、標準モデルを使用するアプリケーションプログラムと標準モデルの仕様との対応を示すアプリケーション仕様対応データベースを前記仕様情報として保持する仕様情報保持手段を備え、前記標準モデル作成手段は、前記仕様情報保持手段に保持されたアプリケーション仕様対応データベースから、起動されるアプリケーションプログラムに対応する仕様を読み出し、読み出した仕様に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

【0062】

これによって、各アプリケーションごとに対応づけられた仕様に沿って標準モデルが作成されるので、アプリケーションごとに最適な標準モデルが作成され、標準モデルが使用される認識システム等における認識精度が向上される。

【0063】

また、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、作成する標準モデルの仕様に関する情報である仕様情報を前記端末装置から受信する仕様情報受信手段を備え、前記標準モデル作成手段は、前記仕様情報受信手段が受信した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

【0064】

これによって、通信路を介して送信されてきた仕様情報に基づいて標準モデルが作成されるので、遠隔制御による標準モデルの生成が可能になるとともに、通信システムを基盤とする認識システムの構築が実現される。

【0065】

たとえば、前記参照モデルおよび前記標準モデルは、1以上のガウス分布を用いて表現され、前記標準モデル作成手段は、前記仕様情報に基づいて、前記標準モデルの混合分布数（ガウス分布の数）を決定してもよい。

【0066】

これによって、作成される標準モデルに含まれるガウス分布の混合分布数が動的に決定されることとなり、認識処理が実行される環境や要求仕様等に応じて標準モデルの構造を制御することが可能となる。例として、標準モデルを使用する認識装置のCPUパワーが小さい場合、記憶容量が小さい場合、要求される認識処理時間が短い場合などは標準モデルの混合分布数を少なく設定して仕様に合わせることができ、一方、要求される認識精度が高い場合などは混合分布数を多く設定して認識精度を高くすることができる。

【0067】

なお、上記利用情報あるいは仕様情報を用いて標準モデルを作成する場合にお

いて、参照モデル準備手段は必ずしも必要ではない。たとえば、利用者の要求に基づいて、あるいは、利用者の要求とは無関係に、予め参照モデルを標準モデル作成装置内に記憶させた状態で標準モデル作成装置を出荷し、利用情報や仕様情報を用いて標準モデルを作成することが可能だからである。

【0068】

また、前記参照モデルおよび前記標準モデルは、1以上のガウス分布を用いて表現され、前記参照モデル記憶手段は、少なくとも1対の参照モデルの混合分布数（ガウス分布の数）が異なる参照モデルを記憶し、前記標準モデル作成手段は、少なくとも1対の参照モデルの混合分布数（ガウス分布の数）が異なる参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

【0069】

これによって、混合分布数が異なる参照モデルに基づいて標準モデルが作成されるので、予め準備された多種多様な構造の参照モデルに基づく標準モデルの作成が可能となり、より認識対象に適した精度の高い標準モデルの作成が実現される。

【0070】

また、前記標準モデル作成装置は、さらに、前記標準モデル作成手段が作成した標準モデルを記憶する標準モデル記憶手段を備えてもよい。

これによって、作成された標準モデルを一時的にバッファリングしておき、送信要求に対してすぐに出力したり、他の装置に提供するデータサーバとしての役割を果たしたりすることが可能となる。

【0071】

また、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、前記標準モデル作成手段が作成した標準モデルを前記端末装置に送信する標準モデル送信手段を備えてもよい。

これによって、作成された標準モデルは空間的に離れた場所に置かれた外部装置に送信されるので、本標準モデル作成装置を標準モデル作成エンジンとして独立させたり、標準モデル作成装置を通信システムにおけるサーバとして機能させた

りすることが可能になる。

【0072】

また、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、前記端末装置から送信される参照モデルを受信する参照モデル受信手段を備え、前記標準モデル作成手段は、少なくとも前記参照モデル受信手段が受信した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

これによって、端末装置が保持した利用環境にふさわしい参照モデルを、通信路を介して送信して、送信した参照モデルを用いて標準モデルを作成できるため、より認識対象に適した精度の高い標準モデルの作成が実現される。例として、利用者Aが環境Aで利用していた参照モデルAが端末装置に保持されており利用者Aは環境Bで利用したい場合、参照モデルAを利用して標準モデルを作成することにより、利用者Aの特徴を反映した精度の高い標準モデルを作成することができる。

【0073】

また、前記参照モデル準備手段は、さらに、前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行ってもよい。たとえば、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、前記端末装置から送信される参照モデルを受信する参照モデル受信手段を備え、前記参照モデル準備手段は、前記参照モデル受信手段が受信した参照モデルを用いて前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行ってもよい。

【0074】

これによって、準備される参照モデルの追加、更新等が行われるので、様々な認識対象用のモデルを参照モデルとして追加したり、より精度の高い参照モデルに置き換えたりすることが可能となり、更新した参照モデルによる標準モデルの再生成や、生成された標準モデルを参照モデルとして再び標準モデルを作成するというフィードバックによる学習等が可能となる。

【0075】

また、前記標準モデル作成手段は、作成する標準モデルの構造を決定する標準モデル構造決定部と、構造が決定された前記標準モデルを特定する統計量の初期値を決定する初期標準モデル作成部と、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を推定して計算する統計量推定部とを有するように構成してもよい。このとき、前記初期標準モデル作成部は、前記統計量推定部が標準モデルの統計量を計算するために用いる、1以上の前記参照モデルを用いて前記標準モデルを特定する統計量の初期値を決定してもよい。たとえば、前記初期標準モデル作成部は、標準モデルの種類を識別するクラスIDに基づいて、前記初期値を決定してもよい。具体的には、前記初期標準モデル作成部は、前記クラスIDと前記初期値と前記参照モデルとの対応を示す対応表を保持し、前記対応表に従って、前記初期値を決定してもよい。

【0076】

これによって、標準モデルが使用される認識の対象の種類ごとにクラスIDを付与しておくことで、最終的に必要とされる標準モデルと共通の性質をもつ初期標準モデルを使用することができるので、精度の高い標準モデルが作成される。

【0077】

なお、本発明は、このような標準モデル作成装置として実現することができるだけでなく、標準モデル作成装置が備える特徴的な構成要素をステップとする標準モデル作成方法として実現したり、それらのステップをコンピュータに実行させるプログラムとして実現したりすることができる。そして、そのプログラムをCD-ROM等の記録媒体やインターネット等の伝送媒体を介して配信することができるのは言うまでもない。

【0078】

【発明の実施の形態】

以下、本発明の実施の形態について図面を参照しながら詳しく説明する。なお、図中同一または相当部分には同一符号を付し、その説明は繰り返さない。

【0079】

(第1の実施の形態)

図1は、本発明の第1の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ101に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデルを作成する場合を例にして説明する。

【0080】

サーバ101は、通信システムにおけるコンピュータ装置等であり、事象の集合と事象または事象間の遷移の出力確率で表現された隠れマルコフモデルによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、読み込み部111と、参照モデル準備部102と、参照モデル記憶部103と、標準モデル作成部104と、書き込み部112とを備える。

【0081】

読み込み部111は、CD-ROMなどのストレージデバイスに書き込まれた子供用参照モデル、成人用参照モデル、高齢者用参照モデルを読み込む。参照モデル準備部102は、読み込まれた参照モデル121を参照モデル記憶部103へ送信する。参照モデル記憶部103は、3個の参照モデル121を記憶する。

【0082】

標準モデル作成部104は、参照モデル記憶部103が記憶した3個 ($N_g=3$) の参照モデル121に対する確率または尤度を最大化または極大化するように標準モデル122を作成する処理部であり、標準モデルの構造 (ガウス分布の混合数など) を決定する標準モデル構造決定部104aと、標準モデルを計算するための統計量の初期値を決定することで初期標準モデルを作成する初期標準モデル作成部104bと、決定された初期標準モデルを記憶する統計量記憶部104cと、統計量記憶部104cに記憶された初期標準モデルに対して、第1近似部104eによる近似計算等を用いることにより、参照モデル記憶部103に記憶されている3個 ($N_g=3$) の参照モデル121に対する確率または尤度を最大化または極大化するような統計量を算出する (最終的な標準モデルを生成する) 統計量推定部104dとからなる。なお、統計量とは、標準モデルを特定するパラメータであり、ここでは、混合重み係数、平均値、分散値である。

【0083】

書き込み部112は、標準モデル作成部104が作成した標準モデル122をCD-ROMなどのストレージデバイスに書き込む。

次に、以上のように構成されたサーバ101の動作について説明する。

【0084】

図2は、サーバ101の動作手順を示すフローチャートである。

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（ステップS100）。つまり、読み込み部111は、CD-ROMなどのストレージデバイスに書き込まれた子供用参照モデル、成人用参照モデル、高齢者用参照モデルを読み込み、参照モデル準備部102は、読み込まれた参照モデル121を参照モデル記憶部103へ送信し、参照モデル記憶部103は、3個の参照モデル121を記憶する。

【0085】

参照モデル121は、音素ごとのHMMにより構成される。参照モデル121の一例を図3に示す。ここでは、子供用参照モデル、成人用参照モデル、高齢者用参照モデルのイメージ図が示されている（なお、本図では、高齢者用参照モデルのイメージ図は省略されている）。これら3個の参照モデルの全てが、状態数3個、各状態は混合分布数が3個の混合ガウス分布によりHMMの出力分布が構成される。特徴量として12次元（ $J=12$ ）のケプストラム係数が用いられる。

【0086】

次に、標準モデル作成部104は、参照モデル記憶部103が記憶した3個の参照モデル121に対する確率または尤度を最大化または極大化するように標準モデル122を作成する（ステップS101）。

【0087】

最後に、書き込み部112は、標準モデル作成部104が作成した標準モデル122をCD-ROMなどのストレージデバイスに書き込む（ステップS102）。CD-ROMなどのストレージデバイスに書き込まれた標準モデルは、子供、成人、高齢者を考慮した音声認識用の標準モデルとして利用される。

【0088】

図4は、図2におけるステップS101（標準モデルの作成）の詳細な手順を示すフローチャートである。

まず、標準モデル構造決定部104aは、標準モデルの構造を決定する（ステップS102a）。ここでは、標準モデルの構造として、音素ごとのHMMにより構成され、3状態であり、各状態における出力分布の混合数を3個（ $M_f=3$ ）と決定する。

【0089】

次に、初期標準モデル作成部104bは、標準モデルを計算するための統計量の初期値を決定する（ステップS102b）。ここでは、参照モデル記憶部103に記憶された3つの参照モデルを、統計処理計算を用いて1つのガウス分布に統合したものを統計量の初期値とし、その初期値を初期標準モデルとして統計量記憶部104cに記憶する。

【0090】

具体的には、初期標準モデル作成部104bは、上記3つの状態I（ $I=1, 2, 3$ ）それぞれについて、以下の数68に示される出力分布を生成する。なお、式中の M_f （ガウス分布の混合数）は、ここでは、3である。

【0091】

【数68】

$$\sum_{m=1}^{M_f} \omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2)$$

【0092】

ここで、

【0093】

【数69】

$$f(x; \mu_{f(m)}, \sigma_{f(m)}^2) \quad (m=1, 2, \dots, M_f)$$

【0094】

は、ガウス分布を表し、

【0 0 9 5】

【数 7 0】

$$x = (x_{(1)}, x_{(2)}, \dots, x_{(J)}) \in R^J$$

【0 0 9 6】

は、1 2 次元 ($J = 1 2$) の L P C ケプストラム係数を表し、

【0 0 9 7】

【数 7 1】

$$\omega_{f(m)} \quad (m = 1, 2, \dots, M_f)$$

【0 0 9 8】

は、各ガウス分布の混合重み係数を表し、

【0 0 9 9】

【数 7 2】

$$\mu_{f(m)} = (\mu_{f(m,1)}, \mu_{f(m,2)}, \dots, \mu_{f(m,J)}) \in R^J \quad (m = 1, 2, \dots, M_f)$$

【0 1 0 0】

は、各ガウス分布の平均値を表し、

【0 1 0 1】

【数 7 3】

$$\sigma_{f(m)}^2 = (\sigma_{f(m,1)}^2, \sigma_{f(m,2)}^2, \dots, \sigma_{f(m,J)}^2) \in R^J \quad (m = 1, 2, \dots, M_f)$$

【0 1 0 2】

は、各ガウス分布の分散値を表す。

そして、統計量推定部 1 0 4 d は、参照モデル記憶部 1 0 3 に記憶された 3 つの参照モデル 1 2 1 を用いて、統計量記憶部 1 0 4 c に記憶された標準モデルの統計量を推定する (ステップ S 1 0 2 c)。

【0 1 0 3】

具体的には、3つ ($N_g = 3$) の参照モデル 1 2 1 の各状態 I ($I = 1, 2, 3$) における出力分布、即ち、以下の数 7 4 に示される出力分布に対する標準モデルの確率または尤度 (以下の数 8 0 に示される尤度 $\log P$) を極大化もしくは最大化するような標準モデルの統計量 (上記数 7 1 に示される混合重み係数、上記数 7 2 に示される平均値、および、上記数 7 3 に示される分散値) を推定する。

【0 1 0 4】

【数 7 4】

$$\sum_{l=1}^{L_{g(i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \quad (i = 1, 2, \dots, N_g)$$

【0 1 0 5】

ここで、

【0 1 0 6】

【数 7 5】

$$g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \quad (i = 1, 2, \dots, N_g, l = 1, 2, \dots, L_{(i)})$$

【0 1 0 7】

はガウス分布を表し、

【0 1 0 8】

【数 7 6】

$$L_{g(i)} \quad (i = 1, 2, \dots, N_g)$$

【0 1 0 9】

は各参照モデルの混合分布数 (ここでは、3) を表し、

【0 1 1 0】

【数 7 7】

$$\nu_{g(i,l)} \quad (l = 1, 2, \dots, L_{g(i)})$$

【0111】

は各ガウス分布の混合重み係数を表し、

【0112】

【数78】

$$\mu_{g(i,l)} \quad (l = 1, 2, \dots, L_{g(i)})$$

【0113】

は各ガウス分布の平均値を表し、

【0114】

【数79】

$$\sigma_{g(i,l)}^2 \quad (l = 1, 2, \dots, L_{g(i)})$$

【0115】

は各ガウス分布の分散値を表す。

【0116】

【数80】

$$\log P = \sum_{l=1}^{N_g} \int_{-\infty}^{\infty} \log \left[\sum_{m=1}^{M_f} \omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2) \right] \left\{ \sum_{l=1}^{L_{g(i)}} v_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx$$

【0117】

そして、以下の数81、数82および数83に示される式に従って、それぞれ、標準モデルの混合重み係数、平均値および分散値を算出する。

【0118】

【数 8 1】

$$\omega_{f(m)} = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{L_{g(i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{k=1}^{M_f} \sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, k) \left\{ \sum_{l=1}^{L_{g(i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f)$$

【0 1 1 9】

【数 8 2】

$$\mu_{f(m,j)} = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) x_{(j)} \left\{ \sum_{l=1}^{L_{g(i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{L_{g(i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0 1 2 0】

【数 8 3】

$$\sigma_{f(m,j)}^2 = \frac{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) (x_{(j)} - \mu_{f(m,j)})^2 \left\{ \sum_{l=1}^{L_{g(i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}{\sum_{i=1}^{N_g} \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{L_{g(i)}} \nu_{g(i,l)} g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \right\} dx}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0 1 2 1】

このとき、統計量推定部 104 d の第 1 近似部 104 e により、以下の数 8 4 に示される近似式が用いられる。

【0 1 2 2】

【数 8 4】

$$\gamma(x, m) = \frac{\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2)}{\sum_{k=1}^{M_f} \omega_{f(k)} f(x, \mu_{f(k)}, \sigma_{f(k)}^2)} \approx \frac{\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2)}{u_{h(m)} h(x, \mu_{h(m)}, \sigma_{h(m)}^2)}$$

$$(m = 1, 2, \dots, M_f)$$

【0 1 2 3】

ここで、

【0 1 2 4】

【数 8 5】

$$u_{h(m)} h(x; \mu_{h(m)}, \sigma_{h(m)}^2) \quad (m = 1, 2, \dots, M_f)$$

【0 1 2 5】

は、

【0 1 2 6】

【数 8 6】

$$u_{h(m)} \quad (m = 1, 2, \dots, M_f)$$

を重みとし、

【0 1 2 7】

【数 8 7】

$$\mu_{h(m)} = (\mu_{h(m,1)}, \mu_{h(m,2)}, \dots, \mu_{h(m,J)}) \in R^J$$

【0 1 2 8】

を平均値とし、

【0 1 2 9】

【数 8 8】

$$\sigma_{h(m)}^2 = (\sigma_{h(m,1)}^2, \sigma_{h(m,2)}^2, \dots, \sigma_{h(m,J)}^2) \in R^J$$

【0 1 3 0】

を分散値とする単一のガウス分布を表す。

また、第 1 近似部 104 e は、上記数 85 に示された単一ガウス分布の重み（数 86）平均値（数 87）および分散値（数 88）を、それぞれ、以下の数 89、数 90 および数 91 に示された式に従って算出する。

【0 1 3 1】

【数 8 9】

$$u_{h(m)} = \sum_{p=1}^{M_f} \omega_{f(m,p)} = \sum_{p=1}^{M_f} \omega_{f(p)} = 1.0 \quad (m = 1, 2, \dots, M_f)$$

【0 1 3 2】

【数 9 0】

$$\mu_{h(m,j)} = \frac{\sum_{p=1}^{M_f} \omega_{f(m,p)} \mu_{f(m,p,j)}}{\sum_{p=1}^{M_f} \omega_{f(m,p)}} = \frac{\sum_{p=1}^{M_f} \omega_{f(p)} \mu_{f(p,j)}}{\sum_{p=1}^{M_f} \omega_{f(p)}}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0 1 3 3】

【数 9 1】

$$\sigma_{h(m,j)}^2 = \frac{\sum_{p=1}^{M_f} \omega_{f(m,p)} (\sigma_{f(m,p,j)}^2 + \mu_{f(m,p,j)}^2)}{\sum_{p=1}^{M_f} \omega_{f(m,p)}} - \mu_{h(m,j)}^2$$

$$= \frac{\sum_{p=1}^{M_f} \omega_{f(p)} (\sigma_{f(p,j)}^2 + \mu_{f(p,j)}^2)}{\sum_{p=1}^{M_f} \omega_{f(p)}} - \mu_{h(m,j)}^2$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0 1 3 4】

図 5 は、第 1 近似部 1 0 4 e による近似計算を説明する図である。第 1 近似部 1 0 4 e は、本図に示されるように、上記数 8 4 に示された近似式における単一ガウス分布（数 8 5）を、標準モデルを構成する全ての混合ガウス分布を用いて決定している。

【0 1 3 5】

以上の第 1 近似部 1 0 4 e による近似式を考慮してまとめると、統計量推定部 1 0 4 d での計算式は次の通りになる。つまり、統計量推定部 1 0 4 d は、以下の数 9 2、数 9 3 および数 9 4 に示される式に従って、それぞれ、混合重み係数、平均値および分散値を算出し、統計量記憶部 1 0 4 c に記憶する。そして、このような統計量の推定と統計量記憶部 1 0 4 c への記憶を R（ ≥ 1 ）回、繰り返す。その結果得られた統計量を最終的に生成する標準モデル 1 2 2 の統計量として出力する。

【0 1 3 6】

【数 9 2】

$$\omega_{f(m)} = \frac{\sum_{i=1}^{N_g} \prod_{j=1}^J \sum_{l=1}^{L_g(i)} A_{(m,l,i,j)}}{\sum_{i=1}^{N_g} \sum_{k=1}^{M_f} \prod_{j=1}^J \sum_{l=1}^{L_g(i)} A_{(k,l,i,j)}} \quad (m=1,2,\dots,M_f)$$

$$A_{(m,l,j)} = \frac{\omega_{f(m)} \sigma_{g(i,j)}^2 \sigma_{h(m,j)}^2}{\sqrt{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 + \sigma_{g(i,j)}^2 \sigma_{h(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(i,j)}^2}}$$

$$\times \exp \left\{ \frac{1}{2} \left(\frac{\left(\frac{\sigma_{f(m,j)} \sigma_{g(i,j)}}{\sigma_{h(m,j)}} \mu_{h(m,j)} - \frac{\sigma_{f(m,j)} \sigma_{h(m,j)}}{\sigma_{g(i,j)}} \mu_{g(i,j)} - \frac{\sigma_{g(i,j)} \sigma_{h(m,j)}}{\sigma_{f(m,j)}} \mu_{f(m,j)} \right)^2}{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 + \sigma_{g(i,j)}^2 \sigma_{h(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(i,j)}^2} + \frac{\mu_{h(m,j)}}{\sigma_{h(m,j)}^2} - \frac{\mu_{g(i,j)}}{\sigma_{g(i,j)}^2} - \frac{\mu_{f(m,j)}}{\sigma_{f(m,j)}^2} \right) \right\}$$

【0 1 3 7】

【数 9 3】

$$\mu_{f(m,j)} = \frac{\sum_{i=1}^{N_g} \sum_{l=1}^{L_g(i)} B_{(m,l,i,j)}}{\sum_{i=1}^{N_g} \prod_{j=1}^J \sum_{l=1}^{L_g(i)} A_{(m,l,i,j)}} \quad (m=1,2,\dots,M_f, j=1,2,\dots,J)$$

$$B_{(m,l,i,j)} = \frac{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 \mu_{g(i,j)} + \sigma_{g(i,j)}^2 \sigma_{h(m,j)}^2 \mu_{f(m,j)} - \sigma_{f(m,j)}^2 \sigma_{g(i,j)}^2 \mu_{h(m,j)}}{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 + \sigma_{g(i,j)}^2 \sigma_{h(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(i,j)}^2} \times A_{(m,l,i,j)}$$

【0 1 3 8】

【数 9 4】

$$\sigma_{f(m,j)}^2 = \frac{\sum_{i=1}^{N_g} \sum_{l=1}^{L_g(i)} C_{(m,j,i,j)}}{\sum_{i=1}^{N_g} \prod_{j=1}^J \sum_{l=1}^{L_g(i)} A_{(m,j,i,j)}} \quad (m = 1, 2, \dots, M_f)$$

$$C_{(m,j,i,j)} = \left\{ \frac{\sigma_{f(m,j)}^2 \sigma_{g(i,j,j)}^2 \sigma_{h(m,j)}^2}{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 + \sigma_{g(i,j,j)}^2 \sigma_{h(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(i,j,j)}^2} + \left(\mu_{f(m,j)} - \frac{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 \mu_{g(i,j,j)} + \sigma_{g(i,j,j)}^2 \sigma_{h(m,j)}^2 \mu_{f(m,j)} - \sigma_{f(m,j)}^2 \sigma_{g(i,j,j)}^2 \mu_{h(m,j)}}{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 + \sigma_{g(i,j,j)}^2 \sigma_{h(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(i,j,j)}^2} \right)^2 \right\} \times A_{(m,j,i,j)}$$

【0139】

なお、状態遷移確率については、HMMの対応する状態遷移確率を参照モデル 121 に対して全て加えあわせた全体が 1 になるように正規化したものを用いる。

【0140】

以上説明したように、本発明の第 1 の実施の形態によれば、予め準備された参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、学習のためのデータや教師データを必要とすることなく簡易に標準モデルが作成されるとともに、既に作成された複数の参照モデルを総合的に勘案した精度の高い標準モデルが作成される。

【0141】

なお、標準モデル 122 は、音素ごとに HMM を構成するに限らず、文脈依存の HMM で構成してもよい。

また、標準モデル作成部 104 は、一部の音素の、一部の状態における事象の出力確率に対してモデル作成を行ってもよい。

【0142】

また、標準モデル 122 を構成する HMM は、音素ごとに異なる状態数により構成してもよいし、状態ごとに異なる分布数の混合ガウス分布により構成してもよい。

【0143】

また、参照モデル121は、子供用参照モデル、成人用参照モデル、高齢者用参照モデルにおいて、異なる状態数により構成してもよいし、異なる混合数の混合ガウス分布により構成してもよい。

【0144】

また、標準モデル122を用いて、サーバ101において音声認識を行ってもよい。

また、参照モデル121をCD-ROM、DVD-RAMなどのストレージデバイスから読み込む代わりに、サーバ101において音声データから参照モデル121を作成してもよい。

【0145】

また、参照モデル準備部102は、必要に応じてCD-ROM、DVD-RAMなどのストレージデバイスから読み込まれた新たな参照モデルを参照モデル記憶部103に追加・更新してもよい。つまり、参照モデル準備部102は、新たな参照モデルを参照モデル記憶部103に格納するだけでなく、同一の認識対象についての参照モデルが参照モデル記憶部103に格納されている場合には、その参照モデルと置き換えることによって参照モデルを更新したり、参照モデル記憶部103に格納されている不要な参照モデルを削除してもよい。

【0146】

また、参照モデル準備部102は、必要に応じて、通信路を介して新たな参照モデルを参照モデル記憶部103に追加・更新してもよい。

また、標準モデルを作成したのちに、さらに音声データにより学習してもよい。

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などのHMMの構造や、状態数などを決定してもよい。

【0147】

(第2の実施の形態)

図6は、本発明の第2の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がセットトップボックス201（以下、STBと呼ぶ）に組み込まれた例が示されている。

本実施の形態では音声認識用の標準モデル（話者適応モデル）を作成する場合を例にして説明する。

【0148】

STB201は、ユーザの発話を認識してTV番組の自動切替等を行うデジタル放送用受信機であり、事象の集合と事象または事象間の遷移の出力確率とによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、マイク211と、音声データ蓄積部212と、参照モデル準備部202と、参照モデル記憶部203と、利用情報作成部204と、参照モデル選択部205と、標準モデル作成部206と、音声認識部213とを備える。

【0149】

マイク211に収集された音声データは、音声データ蓄積部212に蓄積される。参照モデル準備部202は、音声データ蓄積部212が蓄積した音声データを用いて話者ごとに参照モデル221を作成し、参照モデル記憶部203に記憶する。

【0150】

利用情報作成部204は、利用情報224である利用者の音声をマイク211により収集する。ここで、利用情報とは、認識（狭義での認識、識別、認証など）の対象（人・物）に関する情報であり、ここでは、音声認識の対象となる利用者の音声である。参照モデル選択部205は、利用情報作成部204が作成した利用情報224に基づいて、参照モデル記憶部203が記憶している参照モデル221の中から、利用情報224が示す利用者の音声に音響的に近い参照モデル223を選択する。

【0151】

標準モデル作成部206は、参照モデル選択部205が選択した話者の参照モデル223に対する確率または尤度を最大化または極大化するように標準モデル222を作成する処理部であり、標準モデルの構造（ガウス分布の混合分布数など）を決定する標準モデル構造決定部206aと、標準モデルを計算するための統計量の初期値を決定することで初期標準モデルを作成する初期標準モデル作成部206bと、決定された初期標準モデルを記憶する統計量記憶部206cと、

統計量記憶部 206c に記憶された初期標準モデルに対して、一般近似部 206e による近似計算等を用いることにより、参照モデル選択部 205 が選択した参照モデル 223 に対する確率または尤度を最大化または極大化するような統計量を算出する（最終的な標準モデルを生成する）統計量推定部 206d とからなる。

【0152】

音声認識部 213 は、標準モデル作成部 206 によって作成された標準モデル 222 を用いて利用者の音声を認識する。

次に、以上のように構成された STB 201 の動作について説明する。

【0153】

図 7 は、STB 201 の動作手順を示すフローチャートである。

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（ステップ S200）。つまり、マイク 211 により A さんから Z さんの音声データを収集して音声データ蓄積部 212 に蓄積する。参照モデル準備部 202 は、音声データ蓄積部 212 が蓄積した音声データを用いて話者ごとに参照モデル 221 をバウム・ウェルチの再推定の方法により作成する。この処理は、標準モデルの作成が要求される以前に行われる。

【0154】

参照モデル記憶部 203 は、参照モデル準備部 202 が作成した参照モデル 221 を記憶する。参照モデル 221 は、音素ごとの HMM により構成される。参照モデル 221 の一例を図 8 に示す。ここでは、A さんから Z さんの全ての参照モデルが、状態数 3 個、各状態は混合分布数が 5 個の混合ガウス分布により HMM の出力分布が構成される。特徴量として 25 次元 ($J=25$) のメルケプストラム係数が用いられる。

【0155】

ここで、標準モデルの作成が要求される。

利用情報作成部 204 は、利用情報 224 である利用者の音声をマイク 211 により収集する（ステップ S201）。参照モデル選択部 205 は、その利用者の音声に音響的に近い参照モデル 223 を、参照モデル記憶部 203 が記憶して

いる参照モデル221の中から選択する(ステップS202)。具体的には、利用者の音声をAさんからZさんの参照モデルに入力して発声単語に対する尤度が大きい10人($N_g=10$)の話者の参照モデルを選択する。

【0156】

そして、標準モデル作成部206は、参照モデル選択部205が選択した10個の参照モデル223に対する確率または尤度を最大化または極大化するように標準モデル222を作成する(ステップS203)。

【0157】

最後に、音声認識部213は、マイク211から介して送られてくる利用者の音声を入力とし、標準モデル作成部206で作成された標準モデル222を用いて音声認識を行う(S204)。たとえば、利用者が発話した音声を音響解析等を行うことで25次元のメルケプストラム係数を算出し、音素ごとの標準モデル222に入力することで、高い尤度を有する音素の連なりを特定する。そして、その音素の連なりと予め受信している電子番組データ中の番組名とを比較し、一定以上の尤度が検出された場合に、その番組に切り替えるという自動番組切替の制御を行う。

【0158】

次に、図7におけるステップS203(標準モデルの作成)の詳細な手順を説明する。手順の流れは、図4に示されたフローチャートと同様である。ただし、採用する標準モデルの構造や具体的な近似計算等が異なる。

【0159】

まず、標準モデル構造決定部206aは、標準モデルの構造を決定する(図4のステップS102a)。ここでは、標準モデルの構造として、音素ごとのHMMにより構成され、3状態であり、各状態における出力分布の混合分布数が16個($M_f=16$)と決定する。

【0160】

次に、初期標準モデル作成部206bは、標準モデルを計算するための統計量の初期値を決定する(図4のステップS102b)。ここでは、参照モデル選択部205が選択した10個の参照モデル223を、統計処理計算を用いて1つの

ガウス分布に統合したものを統計量の初期値とし、その初期値を初期標準モデルとして統計量記憶部 206c に記憶する。ここでは、話者ごとに学習した混合分布数が 5 の参照モデルを用いて精度の高い混合分布数が 16 (16 混合) の標準モデル (話者適応モデル) を作成する。

【0161】

具体的には、初期標準モデル作成部 206b は、上記 3 つの状態 I (I = 1、2、3) それぞれについて、上記数 68 に示される出力分布を生成する。
ただし、本実施の形態では、上記数 68 に示された出力分布における

【0162】

【数 95】

$$x = (x_{(1)}, x_{(2)}, \dots, x_{(J)}) \in R^J$$

【0163】

は、25 次元 (J = 25) のメルケプストラム係数を表す。

そして、統計量推定部 206d は、参照モデル選択部 205 が選択した 10 個の参照モデル 223 を用いて、統計量記憶部 206c に記憶された標準モデルの統計量を推定する (図 4 のステップ S102c)。

【0164】

つまり、10 個 (Ng = 10) の参照モデル 223 の各状態 I (I = 1、2、3) における出力分布、即ち、上記数 74 に示される出力分布に対する標準モデルの確率 (ここでは、上記数 80 に示される尤度 logP) を極大化もしくは最大化するような標準モデルの統計量 (上記数 71 に示される混合重み係数、上記数 72 に示される平均値、および、上記数 73 に示される分散値) を推定する。

【0165】

ただし、本実施の形態では、上記数 74 に示された出力分布における

【0166】

【数 96】

$$L_{g(i)} \quad (i = 1, 2, \dots, N_g)$$

【0167】

は、5（各参照モデルの混合分布数）である。

具体的には、上記数81、数82および数83に示される式に従って、それぞれ、標準モデルの混合重み係数、平均値および分散値を算出する。

【0168】

このとき、統計量推定部206dの一般近似部206eにより、上記数84に示される近似式が用いられる。

ここで、一般近似部206eは、第1の実施の形態と異なり、上記数84の近似式の分母に示された出力分布

【0169】

【数97】

$$\omega_{f(k)} f(x; \mu_{f(k)}, \sigma_{f(k)}^2) \quad (k=1, 2, \dots, M_f)$$

【0170】

の中から、上記数84の近似式の分子に示された出力分布

【0171】

【数98】

$$\omega_{f(m)} f(x; \mu_{f(m)}, \sigma_{f(m)}^2)$$

【0172】

に距離的に近い3個（ $Ph(m)=3$ ）の出力分布

【0173】

【数99】

$$\omega_{f(m,p)} f(x; \mu_{f(m,p)}, \sigma_{f(m,p)}^2) \quad (m=1, 2, \dots, M_f, p=1, 2, \dots, P_{h(m)})$$

【0174】

を選択し、選択した3個の出力分布を用いて、上記数85に示された単一ガウス分布の重み（数86）、平均値（数87）および分散値（数88）を、それぞれ、以下の数100、数101および数102に示された式に従って算出する。

【0175】

【数100】

$$u_{h(m)} = \sum_{p=1}^{P_{h(m)}} \omega_{f(m,p)} \quad (m=1,2,\dots,M_f)$$

【0176】

【数101】

$$\mu_{h(m,j)} = \frac{\sum_{p=1}^{P_{h(m)}} \omega_{f(m,p)} \mu_{f(m,p,j)}}{\sum_{p=1}^{P_{h(m)}} \omega_{f(m,p)}} \quad (m=1,2,\dots,M_f, j=1,2,\dots,J)$$

【0177】

【数102】

$$\sigma_{h(m,j)}^2 = \frac{\sum_{p=1}^{P_{f(m)}} \omega_{f(m,p)} (\sigma_{f(m,p,j)}^2 + \mu_{f(m,p,j)}^2)}{\sum_{p=1}^{P_{h(m)}} \omega_{f(m,p)}} - \mu_{h(m,j)}^2$$

$$(m=1,2,\dots,M_f, j=1,2,\dots,J)$$

【0178】

図9は、一般近似部206eによる近似計算を説明する図である。一般近似部206eは、本図に示されるように、上記数84に示された近似式における単一ガウス分布(数85)を、標準モデルを構成するMf個の混合ガウス分布の中から、計算対象となる混合ガウス分布に近い一部(P_{h(m)}個)の混合ガウス分布だけを用いて決定している。したがって、全部(Mf個)の混合ガウス分布を用いる第1の実施の形態と比較し、近似計算における計算量が削減される。

【0179】

以上の一般近似部206eによる近似式を考慮してまとめると、統計量推定部

206dでの計算式は次の通りになる。つまり、統計量推定部206dは、以下の数103、数104および数105に示される式に従って、それぞれ、混合重み係数、平均値および分散値を算出し、統計量記憶部206cに記憶する。そして、このような統計量の推定と統計量記憶部206cへの記憶をR(≥1)回、繰り返す。その結果得られた統計量を最終的に生成する標準モデル222の統計量として出力する。なお、繰り返し計算においては、その回数に対応させて、上記近似計算における出力分布の選択個数Ph(m)を小さくし、最終的にPh(m)=1とする計算を行う。

【0180】

【数103】

$$\omega_{f(m)} = \frac{\sum_{l=1}^{N_g} \sum_{i=1}^{L_{g(l)}} \alpha_{(m,l,i)}}{\sum_{k=1}^{M_f} \omega_{f(k)} \left(\sum_{i=1}^{N_g} \sum_{l=1}^{L_{g(l)}} \alpha_{(k,l,i)} \right)} \quad (m=1,2,\dots,M_f)$$

$$\alpha_{(m,l,i)} = v_{g(l,i)} \prod_{j=1}^J D_{(m,l,i,j)}$$

$$D_{(m,l,i,j)} = \frac{\sigma_{H(m,j)}^2}{\sqrt{\sigma_{f(m,j)}^2 \sigma_{H(m,j)}^2 + \sigma_{g(l,j)}^2 \sigma_{H(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(l,j)}^2}} \times \exp \left\{ \frac{1}{2} \left(\frac{\left(\frac{\sigma_{f(m,j)} \sigma_{g(l,j)}}{\sigma_{H(m,j)}} \mu_{H(m,j)} - \frac{\sigma_{f(m,j)} \sigma_{H(m,j)}}{\sigma_{g(l,j)}} \mu_{g(l,j)} - \frac{\sigma_{g(l,j)} \sigma_{H(m,j)}}{\sigma_{f(m,j)}} \mu_{f(m,j)} \right)^2}{\sigma_{f(m,j)}^2 \sigma_{H(m,j)}^2 + \sigma_{g(l,j)}^2 \sigma_{H(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(l,j)}^2} + \frac{\mu_{H(m,j)}}{\sigma_{H(m,j)}^2} \frac{\mu_{g(l,j)}}{\sigma_{g(l,j)}^2} \frac{\mu_{f(m,j)}}{\sigma_{f(m,j)}^2} \right) \right\}$$

【0181】

【数104】

$$\mu_{f(m,j)} = \frac{\sum_{i=1}^{N_g} \sum_{l=1}^{L_{g(l)}} \beta_{(m,l,i,j)} \alpha_{(m,l,i)}}{\sum_{i=1}^{N_g} \sum_{l=1}^{L_{g(l)}} \alpha_{(m,l,i)}} \quad (m=1,2,\dots,M_f, j=1,2,\dots,J)$$

$$\beta_{(m,l,i,j)} = \frac{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 \mu_{g(i,m,j)} + \sigma_{g(i,l,j)}^2 \sigma_{h(m,j)}^2 \mu_{f(m,j)} - \sigma_{f(m,j)}^2 \sigma_{g(i,l,j)}^2 \mu_{h(m,j)}}{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 + \sigma_{g(i,l,j)}^2 \sigma_{h(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(i,l,j)}^2}$$

【0182】

【数105】

$$\sigma_{f(m,j)}^2 = \frac{\sum_{i=1}^{N_g} \sum_{l=1}^{L_{g(l)}} \gamma_{(m,l,i,j)} \alpha_{(m,l,i)}}{\sum_{i=1}^{N_g} \sum_{l=1}^{L_{g(l)}} \alpha_{(m,l,i)}} \quad (m=1,2,\dots,M_f, j=1,2,\dots,J)$$

$$\gamma_{(m,l,i,j)} = \left\{ \frac{\sigma_{f(m,j)}^2 \sigma_{g(i,l,j)}^2 \sigma_{h(m,j)}^2}{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 + \sigma_{g(i,l,j)}^2 \sigma_{h(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(i,l,j)}^2} + \left(\mu_{f(m,j)} - \frac{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 \mu_{g(i,m,j)} + \sigma_{g(i,l,j)}^2 \sigma_{h(m,j)}^2 \mu_{f(m,j)} - \sigma_{f(m,j)}^2 \sigma_{g(i,l,j)}^2 \mu_{h(m,j)}}{\sigma_{f(m,j)}^2 \sigma_{h(m,j)}^2 + \sigma_{g(i,l,j)}^2 \sigma_{h(m,j)}^2 - \sigma_{f(m,j)}^2 \sigma_{g(i,l,j)}^2} \right)^2 \right\}$$

【0183】

なお、状態遷移確率については、HMMの対応する状態遷移確率を参照モデル223に対して全て加えあわせた全体が1になるように正規化したものを用いる。

【0184】

以上説明したように、本発明の第2の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況

によりふさわしい精度の高い標準モデルが提供される。

【0185】

なお、標準モデル222は、音素ごとにHMMを構成するに限らず、文脈依存のHMMで構成してもよい。

また、標準モデル作成部206は、一部の音素の、一部の状態における事象の出力確率に対してモデル作成を行ってもよい。

【0186】

また、標準モデル222を構成するHMMは、音素ごとに異なる状態数により構成してもよいし、状態ごとに異なる分布数の混合ガウス分布により構成してもよい。

【0187】

また、参照モデル221は、話者ごとHMMにおいて、異なる状態数により構成してもよいし、異なる混合数の混合ガウス分布により構成してもよい。

また、参照モデル221は、話者ごとHMMに限らず、話者・雑音・声の調子ごとに作成してもよい。

【0188】

また、標準モデル222をCD-ROM、ハードディスク、DVD-RAMなどのストレージデバイスに記録してもよい。

また、参照モデル221を作成する代わりに、CD-ROM、DVD-RAMなどのストレージデバイスから読み込んでもよい。

【0189】

また、参照モデル選択部205は、利用情報224に基づいて利用者ごとに選択する参照モデルの数を変えてもよい。

また、参照モデル準備部202は、必要に応じて新たな参照モデルを作成して参照モデル記憶部203に追加・更新してもよいし、参照モデル記憶部203に格納されている不要な参照モデルを削除してもよい。

【0190】

また、参照モデル準備部202は、必要に応じて、通信路を介して新たな参照モデルを参照モデル記憶部203に追加・更新してもよい。

また、上記近似計算において選択する出力分布の個数 $Ph(m)$ は、対象とする事象や標準モデルの出力分布によって異なってもよいし、分布間距離に基づいて決定してもよい。

【0191】

また、標準モデルを作成したのちに、さらに音声データにより学習してもよい。

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などのHMMの構造や、状態数などを決定してもよい。

【0192】

(第3の実施の形態)

図10は、本発明の第3の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がPDA(Personal Digital Assistant)301に組み込まれた例が示されている。本実施の形態では雑音識別用の標準モデル(雑音モデル)を作成する場合を例にして説明する。

【0193】

PDA301は、携帯情報端末であり、事象の出力確率によって定義される雑音識別用の標準モデルを作成する標準モデル作成装置として、読み込み部311と、参照モデル準備部302と、参照モデル記憶部303と、利用情報作成部304と、参照モデル選択部305と、標準モデル作成部306と、仕様情報作成部307と、マイク312と、雑音識別部313とを備える。

【0194】

読み込み部311は、CD-ROMなどのストレージデバイスに書き込まれた乗用車Aの参照モデル、乗用車Bの参照モデル、バスAの参照モデル、小雨の参照モデル、大雨の参照モデルなどの雑音の参照モデルを読み込む。参照モデル準備部302は、読み込まれた参照モデル321を参照モデル記憶部303へ送信する。参照モデル記憶部303は、参照モデル321を記憶する。

【0195】

利用情報作成部304は、利用情報324である雑音の種類をPDA301の

画面とキーを利用して作成する。参照モデル選択部 305 は、利用情報 324 である雑音の種類に音響的に近い参照モデルを、参照モデル記憶部 303 が記憶している参照モデル 321 の中から選択する。仕様情報作成部 307 は、PDA 301 の仕様に基づき仕様情報 325 を作成する。ここで、仕様情報とは、作成する標準モデルの仕様に関する情報であり、ここでは、PDA 301 が備える CPU の処理能力に関する情報である。

【0196】

標準モデル作成部 306 は、仕様情報作成部 307 で作成された仕様情報 325 に基づいて、参照モデル選択部 305 が選択した雑音の参照モデル 323 に対する確率または尤度を最大化または極大化するように標準モデル 322 を作成する処理部であり、標準モデルの構造（ガウス分布の混合分布数など）を決定する標準モデル構造決定部 306a と、標準モデルを計算するための統計量の初期値を決定することで初期標準モデルを作成する初期標準モデル作成部 306b と、決定された初期標準モデルを記憶する統計量記憶部 306c と、統計量記憶部 306c に記憶された初期標準モデルに対して、第 2 近似部 306e による近似計算等を用いることにより、参照モデル選択部 305 が選択した参照モデル 323 に対する確率または尤度を最大化または極大化するような統計量を算出する（最終的な標準モデルを生成する）統計量推定部 306d とからなる。

【0197】

雑音識別部 313 は、標準モデル作成部 306 で作成された標準モデル 322 を用いて、マイク 312 から入力された雑音の種類を識別する。

次に、以上のように構成された PDA 301 の動作について説明する。

【0198】

図 11 は、PDA 301 の動作手順を示すフローチャートである。

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（ステップ S300）。つまり、読み込み部 311 は、ストレージデバイスに書き込まれた雑音の参照モデルを読み込み、参照モデル準備部 302 は、読み込まれた参照モデル 321 を参照モデル記憶部 303 へ送信し、参照モデル記憶部 303 は、参照モデル 321 を記憶する。

【0199】

参照モデル321は、GMMより構成される。参照モデル321の一例を図12に示す。ここでは、各雑音モデルは混合分布数が3個のGMMにより構成される。特徴量として5次元 ($J=5$) のLPCケプストラム係数が用いられる。

【0200】

次に、利用情報作成部304は、識別したい雑音の種類である利用情報324を作成する(ステップS301)。図13にPDA301の選択画面の一例を示す。ここでは、乗用車の雑音が選択される。参照モデル選択部305は、選択された利用情報324である乗用車の雑音に音響的に近い参照モデルである乗用車Aの参照モデルと乗用車Bの参照モデルを、参照モデル記憶部303が記憶している参照モデル321の中から選択する(ステップS302)。

【0201】

そして、仕様情報作成部307は、PDA301の仕様に基づき、仕様情報325を作成する(ステップS303)。ここでは、PDA301のCPUの仕様に基づきCPUパワーが小さいという仕様情報325を作成する。標準モデル作成部306は、作成された仕様情報325に基づいて、参照モデル選択部305が選択した参照モデル323に対する確率または尤度を最大化または極大化するように標準モデル322を作成する(ステップS304)。

【0202】

最後に、雑音識別部313は、利用者によってマイク312から入力された雑音に対して、標準モデル322を用いて、雑音の識別を行う(ステップS305)。

【0203】

次に、図11におけるステップS304(標準モデルの作成)の詳細な手順を説明する。手順の流れは、図4に示されたフローチャートと同様である。ただし、採用する標準モデルの構造や具体的な近似計算等が異なる。

【0204】

まず、標準モデル構造決定部306aは、標準モデルの構造を決定する(図4のステップS102a)。ここでは、標準モデルの構造として、仕様情報325

であるCPUパワーが小さいという情報に基づいて1混合 ($M_f = 1$) のGMMにより標準モデル322を構成すると決定する。

【0205】

次に、初期標準モデル作成部306bは、標準モデルを計算するための統計量の初期値を決定する(図4のステップS102b)。ここでは、選択された参照モデル323である乗用車Aの3混合の参照モデルを、統計処理計算を用いて1つのガウス分布に統合したものを統計量の初期値として統計量記憶部306cに記憶する。

【0206】

具体的には、初期標準モデル作成部306bは、上記数68に示される出力分布を生成する。

ただし、本実施の形態では、上記数68に示された出力分布における

【0207】

【数106】

$$x = (x_{(1)}, x_{(2)}, \dots, x_{(J)}) \in R^J$$

【0208】

は、5次元 ($J = 5$) のLPCケプストラム係数を表す。

そして、統計量推定部306dは、参照モデル選択部305が選択した2個の参照モデル323を用いて、統計量記憶部306cに記憶された標準モデルの統計量を推定する(図4のステップS102c)。

【0209】

つまり、2個 ($N_g = 2$) の参照モデル323における出力分布、即ち、上記数74に示される出力分布に対する標準モデルの確率(ここでは、上記数80に示される尤度 $\log P$)を極大化もしくは最大化するような標準モデルの統計量(上記数71に示される混合重み係数、上記数72に示される平均値、および、上記数73に示される分散値)を推定する。

【0210】

ただし、本実施の形態では、上記数74に示された出力分布における

【0211】

【数107】

$$L_{g(i)} \quad (i=1,2,\dots,N_g)$$

【0212】

は、3（各参照モデルの混合分布数）である。

具体的には、上記数81、数82および数83に示される式に従って、それぞれ、標準モデルの混合重み係数、平均値および分散値を算出する。

【0213】

このとき、統計量推定部306dの第2近似部306eは、標準モデルの各ガウス分布はお互いに影響を与えないと仮定して、以下の近似式を用いる。

【0214】

【数108】

$$\gamma(x, m) \approx \frac{\omega_{f(m)} f(x, \mu_{f(m)}, \sigma_{f(m)}^2)}{u_{h(m)} h(x, \mu_{h(m)}, \sigma_{h(m)}^2)} \approx 1.0$$

$$(m=1,2,\dots,M_f)$$

【0215】

また、標準モデルのガウス分布

【0216】

【数109】

$$\omega_{f(m,p)} f(x, \mu_{f(m,p)}, \sigma_{f(m,p)}^2) \quad (m=1,2,\dots,M_f, p=1,2,\dots,P_{h(m)})$$

【0217】

の近傍の

【0218】

【数 1 1 0】

 x

【0 2 1 9】

とは、前記数 1 0 9 が示す出力分布との平均値のユークリッド距離、マハラノビス距離、カルバック・ライブラー (KL) 距離などの分布間距離に近い $Qg(m, i)$ 個の参照モデル 3 2 3 のガウス分布

【0 2 2 0】

【数 1 1 1】

$$g(x; \mu_{g(i,l)}, \sigma_{g(i,l)}^2) \quad (i = 1, 2, \dots, N_g, l = 1, 2, \dots, L_{(i)})$$

【0 2 2 1】

が存在する空間であって、

【0 2 2 2】

【数 1 1 2】

$$\omega_{f(m,p)} f(x; \mu_{f(m,p)}, \sigma_{f(m,p)}^2) \quad (m = 1, 2, \dots, M_f, p = 1, 2, \dots, P_{h(m)})$$

【0 2 2 3】

との分布間距離に近い $Qg(m, i)$ 個 ($1 \leq Qg(m, i) \leq Lg(i)$) の前記参照ベクトルの出力分布とは、前記参照モデルの出力分布

【0 2 2 4】

【数 1 1 3】

$$v_{g(i,l)} g(x; \mu_{g(l)}, \sigma_{g(l)}^2) \quad (i = 1, 2, \dots, N_g, l = 1, 2, \dots, L_{g(i)})$$

【0 2 2 5】

のうち分布間距離が 1 番近い (近傍指示パラメータ $G = 1$) 前記標準モデルの出力分布が前記数 5 2 である前記参照ベクトルの出力分布であると近似する。

【0 2 2 6】

図 1 4 は、この統計量推定部 3 0 6 d による統計量の推定手順を示す概念図で

ある。各参照モデルの各ガウス分布に対して、平均値のユークリッド距離、マハラノビス距離などの分布間距離が最も近いものが標準モデルのガウス分布 m であるガウス分布を用いて統計量の推定を行うことが示されている。

【0227】

図15は、第2近似部306eによる近似計算を説明する図である。第2近似部306eは、本図に示されるように、各参照モデルの各ガウス分布に対して、距離が最も近い標準モデルのガウス分布 m を決定することで、上記数108に示された近似式を用いている。

【0228】

以上の第2近似部306eによる近似式を考慮してまとめると、統計量推定部306dでの計算式は次の通りになる。つまり、統計量推定部306dは、以下の数114、数115および数116に示される式に従って、それぞれ、混合重み係数、平均値および分散値を算出し、それらのパラメータによって特定される標準モデルを最終的な標準モデル322として生成する。

【0229】

【数114】

$$\omega_{f(m)} = \frac{\sum_{i=1}^{N_g} \sum_{l=1}^{Q_{g(m,l)}} v_{g(i,l)}}{\sum_{k=1}^{M_f} \sum_{i=1}^{N_g} \sum_{l=1}^{Q_{g(m,l)}} v_{g(i,l)}} \quad (m = 1, 2, \dots, M_f)$$

【0230】

(ここで、分母、分子の和は、各参照モデルの各ガウス分布に対して、平均値のユークリッド距離、マハラノビス距離などの分布間距離が最も近いものが標準モデルのガウス分布 m であるガウス分布に関する和を意味する。)

【0231】

【数115】

$$\mu_{f(m,j)} = \frac{\sum_{i=1}^{N_g} \sum_{l=1}^{Q_{g(m,l)}} \nu_{g(i,l)} \mu_{g(i,l,j)}}{\sum_{i=1}^{N_g} \sum_{l=1}^{Q_{g(m,l)}} \nu_{g(i,l)}}$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0232】

(ここで、分母、分子の和は、各参照モデルの各ガウス分布に対して、平均値のユークリッド距離、マハラノビス距離などの分布間距離が最も近いものが標準モデルのガウス分布 m であるガウス分布に関する和を意味する。)

【0233】

【数116】

$$\sigma_{f(m,j)}^2 = \frac{\sum_{i=1}^{N_g} \sum_{l=1}^{Q_{g(m,l)}} \nu_{g(i,l)} (\sigma_{g(i,l)}^2 + \mu_{g(i,l,j)}^2)}{\sum_{i=1}^{N_g} \sum_{l=1}^{Q_{g(m,l)}} \nu_{g(i,l)}} - \mu_{f(m,j)}^2$$

$$(m = 1, 2, \dots, M_f, j = 1, 2, \dots, J)$$

【0234】

(ここで、分母、分子の和は、各参照モデルの各ガウス分布に対して、平均値のユークリッド距離、マハラノビス距離などの分布間距離が最も近いものが標準モデルのガウス分布 m であるガウス分布に関する和を意味する。)

ただし、

【0235】

【数 117】

$$\sum_{i=1}^{N_g} Q_{g(m,i)} = 0 \quad (m=1,2,\dots,M_f)$$

【0236】

の場合において、

(第1の方法) 混合重み係数、平均値、分散値を更新しない。

(第2の方法) 混合重み係数の値をゼロにして、平均値、分散値を所定の値にする。

(第3の方法) 混合重み係数の値を所定の値にして、平均値、分散値を標準モデルの出力分布を1個の分布に表現したときの平均値、分散値にする。

【0237】

のいずれかを利用して統計量の値を決定する。なお、利用する方法は、繰り返し回数R、HMM、HMMの状態ごとに異なってもよい。ここでは、第1の方法を用いる。

【0238】

統計量推定部306dは、このように推定した標準モデルの統計量を統計量記憶部306cに記憶する。そして、このような統計量の推定と統計量記憶部306cへの記憶をR(≧1)回、繰り返す。その結果得られた統計量を最終的に生成する標準モデル322の統計量として出力する。

【0239】

以上説明したように、本発明の第3の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい精度の高い標準モデルが提供される。

【0240】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

なお、統計量推定部306dによる処理の繰り返し回数は、上記数80に示さ

れた尤度の大きさがある一定のしきい値以上になるまでの回数としてもよい。

【0241】

また、標準モデル322を構成するGMMは、雑音の種類ごとに異なる混合分布数の混合ガウス分布により構成してもよい。

また、識別モデルは、雑音モデルに限らず、話者を識別してもよいし、年齢などを識別してもよい。

【0242】

また、標準モデル322をCD-ROM、DVD-RAM、ハードディスクなどのストレージデバイスに記録してもよい。

また、参照モデル321をCD-ROMなどのストレージデバイスから読み込む代わりに、PDA301において雑音データから参照モデル321を作成してもよい。

【0243】

また、参照モデル準備部302は、必要に応じてCD-ROMなどのストレージデバイスから読み込まれた新たな参照モデルを参照モデル記憶部303に追加・更新してもよいし、参照モデル記憶部303に格納されている不要な参照モデルを削除してもよい。

【0244】

また、参照モデル準備部302は、必要に応じて、通信路を介して新たな参照モデルを参照モデル記憶部303に追加・更新してもよい。

また、標準モデルを作成したのちに、さらにデータにより学習してもよい。

【0245】

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などのHMMの構造や、状態数などを決定してもよい。

また、近傍指示パラメータGは、対象とする事象や標準モデルの出力分布によって異なってもよいし、繰り返し回数Rによって変化させてもよい。

【0246】

(第4の実施の形態)

図16は、本発明の第4の実施の形態における標準モデル作成装置の全体構成

を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ401に組み込まれた例が示されている。本実施の形態では顔認識用の標準モデルを作成する場合を例にして説明する。

【0247】

サーバ401は、通信システムにおけるコンピュータ装置等であり、事象の出力確率によって定義される顔認識用の標準モデルを作成する標準モデル作成装置として、カメラ411と、画像データ蓄積部412と、参照モデル準備部402と、参照モデル記憶部403と、利用情報受信部404と、参照モデル選択部405と、標準モデル作成部406と、書き込み部413とを備える。

【0248】

カメラ411により、顔の画像データが収集され、画像データ蓄積部412に顔画像データが蓄積される。参照モデル準備部402は、画像データ蓄積部412が蓄積した顔画像データを用いて話者ごとに参照モデル421を作成し、参照モデル記憶部403に記憶する。

【0249】

利用情報受信部404は、利用者が希望する顔認識の対象となる人間の年齢の年代と性別の情報を利用情報424として電話414により受信する。参照モデル選択部405は、利用情報受信部404が受信した利用情報424に基づいて、参照モデル記憶部403が記憶している参照モデル421の中から、利用情報424が示す年代と性別の話者に対応する参照モデル423を選択する。

【0250】

標準モデル作成部406は、参照モデル選択部405が選択した話者の顔画像の参照モデル423に対する確率または尤度を最大化または極大化するように標準モデル422を作成する処理部であり、第2の実施の形態における標準モデル作成部206と同一の機能を有するとともに、第1の実施の形態における第1近似部104eと第3の実施の形態における第2近似部306eの機能を有する。つまり、第1～第3の実施の形態で示された3種類の近似計算を組み合わせた計算を行う。

【0251】

書き込み部 413 は、標準モデル作成部 406 が作成した標準モデル 422 を CD-ROM などのストレージデバイスに書き込む。

次に、以上のように構成されたサーバ 401 の動作について説明する。

【0252】

図 17 は、サーバ 401 の動作手順を示すフローチャートである。図 18 は、サーバ 401 の動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【0253】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図 17 のステップ S400）。つまり、カメラ 411 により A さんから Z さんの顔画像データを収集して画像データ蓄積部 412 に蓄積する。参照モデル準備部 402 は、画像データ蓄積部 412 が蓄積した顔画像データを用いて、話者ごとの参照モデル 421 を EM アルゴリズムにより作成する。ここでは参照モデル 421 は GMM で構成される。

【0254】

参照モデル記憶部 403 は、参照モデル準備部 402 が作成した参照モデル 421 を記憶する。ここでは、図 18 の参照モデル 421 に示されるように、A さんから Z さんの全ての参照モデルが、混合分布数が 5 個の GMM により構成される。特徴量として 100 次元（ $J = 100$ ）の画素の濃度値を用いる。

【0255】

次に、利用情報受信部 404 は、利用情報 424 である年代と性別の情報を電話 414 により受信する（図 17 のステップ S401）。ここでは、利用情報 424 として、11 歳から 15 歳の男性と 22 歳から 26 歳の女性である。参照モデル選択部 405 は、その利用情報 424 に基づいて、参照モデル記憶部 403 が記憶している参照モデル 421 から、利用情報 424 に対応する参照モデル 423 を選択する（図 17 のステップ S402）。具体的には、図 18 の「選択された参照モデル 423」に示されるように、ここでは、11 歳から 15 歳の男性および 22 歳から 26 歳の女性の参照モデルを選択する。

【0256】

そして、標準モデル作成部 406 は、参照モデル選択部 405 が選択した話者の参照モデル 423 に対する確率または尤度を最大化または極大化するように標準モデル 422 を作成する（図 17 のステップ S403）。ここでは、図 18 の標準モデル 422 に示されるように、2 つの標準モデル 422 それぞれを、混合分布数が 3 個の GMM により構成する。

【0257】

標準モデル 422 の作成方法は、基本的には、第 2 の実施の形態と同様に行われる。ただし、標準モデル 422 の統計量の推定における近似計算については、具体体には、以下のようにして行われる。つまり、標準モデル作成部 406 は、内蔵の記憶部等を介することで、第 1 の実施の形態における第 1 近似部 104 e による近似計算と同様の近似計算によって作成したモデルを初期値として、第 2 の実施の形態における一般近似部 206 e による近似計算と同様の近似計算による計算を行い、その結果を初期値として第 3 の実施の形態における第 2 近似部 306 e による近似計算と同様の近似計算を行う。

【0258】

書き込み部 413 は、標準モデル作成部 406 が作成した 2 つの標準モデル 422 を CD-ROM などのストレージデバイスに書き込む（図 17 のステップ S404）。

【0259】

利用者は、11 歳から 15 歳の男性の標準モデルと 22 歳から 26 歳の女性の標準モデルが書き込まれたストレージデバイスを郵送で受け取る。

以上説明したように、本発明の第 4 の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい高精度な標準モデルが提供される。

【0260】

なお、標準モデル 422 を構成する GMM は、話者ごとに異なる分布数の混合ガウス分布により構成してもよい。

また、参照モデル準備部 402 は、必要に応じて新たな参照モデルを作成して

参照モデル記憶部 403 に追加・更新してもよいし、参照モデル記憶部 403 に格納されている不要な参照モデルを削除してもよい。

【0261】

また、標準モデルを作成したのちに、さらにデータにより学習してもよい。

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などの HMM の構造や、状態数などを決定してもよい。

【0262】

(第 5 の実施の形態)

図 19 は、本発明の第 5 の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ 501 に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデル（適応モデル）を作成する場合を例にして説明する。

【0263】

サーバ 501 は、通信システムにおけるコンピュータ装置等であり、事象の集合と事象または事象間の遷移の出力確率とによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、読み込み部 511 と、音声データ蓄積部 512 と、参照モデル準備部 502 と、参照モデル記憶部 503 と、利用情報受信部 504 と、参照モデル選択部 505 と、標準モデル作成部 506 と、仕様情報受信部 507 と、書き込み部 513 とを備える。

【0264】

読み込み部 511 は、CD-ROM などのストレージデバイスに書き込まれた子供、成人、高齢者の音声データを読み込み、音声データ蓄積部 512 に蓄積する。参照モデル準備部 502 は、音声データ蓄積部 512 が蓄積した音声データを用いて話者ごとに参照モデル 521 を作成する。参照モデル記憶部 503 は、参照モデル準備部 502 が作成した参照モデル 521 を記憶する。

【0265】

仕様情報受信部 507 は、仕様情報 525 を受信する。利用情報受信部 504 は、利用情報 524 である利用者の音声を受信する。参照モデル選択部 505 は

、利用情報 524 である利用者の音声に音響的に近い話者の参照モデルを、参照モデル記憶部 503 が記憶している参照モデル 521 から選択する。

【0266】

標準モデル作成部 506 は、仕様情報 525 に基づいて、参照モデル選択部 505 が選択した話者の参照モデル 523 に対する確率または尤度を最大化または極大化するように標準モデル 522 を作成する処理部であり、第 1 の実施の形態における標準モデル作成部 104 と同一の機能を有する。書き込み部 513 は、標準モデル作成部 506 が作成した標準モデル 522 を CD-ROM などのストレージデバイスに書き込む。

【0267】

次に、以上のように構成されたサーバ 501 の動作について説明する。

図 20 は、サーバ 501 の動作手順を示すフローチャートである。図 21 は、サーバ 501 の動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【0268】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図 20 のステップ S500）。つまり、読み込み部 511 は、CD-ROM などのストレージデバイスに書き込まれた音声データを読み込み、音声データ蓄積部 512 に蓄積する。参照モデル準備部 502 は、音声データ蓄積部 512 が蓄積した音声データを用いて話者ごとに参照モデル 521 をバウム・ウェルチの再推定の方法により作成する。参照モデル記憶部 503 は、参照モデル準備部 502 が作成した参照モデル 521 を記憶する。

【0269】

参照モデル 521 は、音素ごとの HMM により構成される。ここでは、図 21 の参照モデル 521 に示されるように、子供の各話者の参照モデルは、状態数 3 個、各状態は混合分布数が 3 個の混合ガウス分布により HMM の出力分布が構成され、成人の各話者の参照モデルが、状態数 3 個、各状態は混合分布数が 64 個の混合ガウス分布により HMM の出力分布が構成され、高齢者の各話者の参照モデルは、状態数 3 個、各状態は混合分布数が 16 個の混合ガウス分布により HM

Mの出力分布が構成される。これは、子供の音声データが比較的少なく、成人の音声データが多いためである。特徴量として25次元 ($J=25$) のメルケプストラム係数が用いられる。

【0270】

次に、利用情報受信部504は、利用者の音声を、端末装置514から、利用情報524として受信する(図20のステップS501)。参照モデル選択部505は、利用情報524である利用者の音声に音響的に近い参照モデル523を、参照モデル記憶部503が記憶している参照モデル521から選択する(図20のステップS502)。具体的には、図21の「選択された参照モデル523」に示されるように、ここでは、近い話者10人($N_g=10$)の参照モデルが選択される

そして、仕様情報受信部507は、利用者の要求に基づき仕様情報525を端末装置514から受信する(図20のステップS503)。ここでは、速い認識処理という仕様情報525を受信する。標準モデル作成部506は、仕様情報受信部507が受信した仕様情報525に基づいて、参照モデル選択部505が選択した話者の参照モデル523に対する確率または尤度を最大化または極大化するように標準モデル522を作成する(図20のステップS504)。具体的には、標準モデル522は、図21の標準モデル522に示されるように、仕様情報525である速い認識処理という情報に基づいて、2混合($M_f=2$)で、3状態のHMMより構成する。HMMは音素ごとに構成する。

【0271】

標準モデル522の作成方法は、第1の実施の形態と同様に行われる。

書き込み部513は、標準モデル作成部506が作成した標準モデル522をCD-ROMなどのストレージデバイスに書き込む(図20のステップS505)。

【0272】

以上説明したように、本発明の第5の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況

によりふさわしい精度の高い標準モデルが提供される。

【0273】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

また、参照モデル準備部 502 において、参照モデルごとにデータ数に適した混合分布数の精度の高い参照モデルを準備でき、精度の高い参照モデルを用いて標準モデルを作成できる。このため精度の高い標準モデルの利用が可能となる。

【0274】

なお、標準モデル 522 は、音素ごとに HMM を構成するに限らず、文脈依存の HMM で構成してもよい。

また、標準モデル 522 を構成する HMM は、状態ごとに異なる分布数の混合ガウス分布により構成してもよい。

【0275】

また、標準モデル 522 を用いて、サーバ 501 において音声認識を行ってもよい。

また、参照モデル準備部 502 は、必要に応じて新たな参照モデルを作成して参照モデル記憶部 503 に追加・更新してもよいし、参照モデル記憶部 503 に格納されている不要な参照モデルを削除してもよい。

【0276】

また、標準モデルを作成したのちに、さらにデータにより学習してもよい。

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などの HMM の構造や、状態数などを決定してもよい。

【0277】

(第 6 の実施の形態)

図 22 は、本発明の第 6 の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ 601 に組み込まれた例が示されている。本実施の形態では意図理解のための標準モデル（嗜好モデル）を作成する場合を例にして説明する。

【0278】

サーバ601は、通信システムにおけるコンピュータ装置等であり、事象の出力確率によって定義される意図理解用の標準モデルを作成する標準モデル作成装置として、読み込み部611と、参照モデル準備部602と、参照モデル記憶部603と、利用情報受信部604と、参照モデル選択部605と、標準モデル作成部606と、仕様情報作成部607とを備える。

【0279】

読み込み部611は、CD-ROMなどのストレージデバイスに書き込まれた年齢別の話者Aさんから話者Zさんの嗜好モデルを読み込み、参照モデル準備部602は、読み込まれた参照モデル621を参照モデル記憶部603へ送信し、参照モデル記憶部603は、参照モデル621を記憶する。

【0280】

仕様情報作成部607は、普及しているコンピュータのCPUパワーに合わせて仕様情報625を作成する。利用情報受信部604は、端末装置614から利用情報624を受信する。参照モデル選択部605は、利用情報受信部604が受信した利用情報624に基づいて、参照モデル記憶部603が記憶している参照モデル621からの中から、利用情報624に対応した参照モデル623を選択する。

【0281】

標準モデル作成部606は、仕様情報作成部607が作成した仕様情報625に基づいて、参照モデル選択部605が選択した参照モデル623に対する確率または尤度を最大化または極大化するように標準モデル622を作成する処理部であり、第2の実施の形態における標準モデル作成部206と同一の機能を有するとともに、第3の実施の形態における第2近似部306eの機能を有する。つまり、第2および第3の実施の形態で示された2種類の近似計算を組み合わせた計算を行う。

【0282】

次に、以上のように構成されたサーバ601の動作について説明する。

図23は、サーバ601の動作手順を示すフローチャートである。図24は、

サーバ601の動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【0283】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図23のステップS600）。つまり、読み込み部611は、CD-ROMなどのストレージデバイスに書き込まれた年齢別の話者Aさんから話者Zさんの嗜好モデルを読み込み、参照モデル準備部602は、読み込まれた参照モデル621を参照モデル記憶部603へ送信し、参照モデル記憶部603は、参照モデル621を記憶する。

【0284】

参照モデル621は、GMMより構成される。ここでは、図24の参照モデル621に示されるように、混合分布数が3個のGMMにより構成される。学習データとして、趣味、性格などを数値化した5次元（ $J=5$ ）の特徴量を用いる。参照モデルの準備は、標準モデルの作成が要求される以前に行う。

【0285】

次に、利用情報受信部604は、嗜好モデルを作成したい年齢層である利用情報624を受信する（図23のステップS601）。ここでは、20代、30代、40代の年代別の嗜好モデルを利用するという利用情報624である。参照モデル選択部605は、図24の「選択された参照モデル623」に示されるように、利用情報受信部604が受信した利用情報624が示す年代の話者の嗜好モデルを、参照モデル記憶部603が記憶している参照モデル621から選択する（図23のステップS602）。

【0286】

そして、仕様情報作成部607は、普及しているコンピュータのCPUパワー、記憶容量などに基づき仕様情報625を作成する（図23のステップS603）。ここでは、通常速度の認識処理という仕様情報625を作成する。

【0287】

標準モデル作成部606は、仕様情報作成部607が作成した仕様情報625に基づいて、参照モデル選択部605が選択した話者の参照モデル623に対す

る確率または尤度を最大化または極大化するように標準モデル 622 を作成する (図 23 のステップ S604)。ここでは、標準モデル 622 は、図 24 の標準モデル 622 に示されるように、仕様情報 625 である通常速度の認識処理という情報に基づいて 3 混合 ($Mf=3$) の GMM より構成する。

【0288】

標準モデル 622 の作成方法は、基本的には、第 2 の実施の形態と同様に行われる。ただし、標準モデル 622 の統計量の推定における近似計算については、具体体には、以下のようにして行われる。つまり、標準モデル作成部 606 は、内蔵の記憶部等を介することで、第 2 の実施の形態における一般近似部 206e による近似計算と同様の近似計算による計算を行い、その結果を初期値として第 3 の実施の形態における第 2 近似部 306e による近似計算と同様の近似計算を行う。

【0289】

以上説明したように、本発明の第 6 の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい精度の高い標準モデルが提供される。

【0290】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

なお、標準モデル 622 を構成する GMM は、話者ごとに異なる分布数の混合ガウス分布により構成してもよい。

【0291】

また、参照モデル準備部 602 は、必要に応じて CD-ROM などのストレージデバイスから読み込まれた新たな参照モデルを参照モデル記憶部 603 に追加・更新してもよいし、参照モデル記憶部 603 に格納されている不要な参照モデルを削除してもよい。

【0292】

また、参照モデルおよび標準モデルの GMM はベイジアンネットの一部を表現

するものでもよい。

また、標準モデルを作成したのちに、さらにデータにより学習してもよい。

【0293】

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などのHMMの構造や、状態数などを決定してもよい。

【0294】

(第7の実施の形態)

図25は、本発明の第7の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ701に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデル（適応モデル）を作成する場合を例にして説明する。

【0295】

サーバ701は、通信システムにおけるコンピュータ装置等であり、事象の集合と事象または事象間の遷移の出力確率とによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、読み込み部711と、参照モデル準備部702と、参照モデル記憶部703と、利用情報受信部704と、参照モデル選択部705と、標準モデル作成部706と、仕様情報受信部707と、標準モデル記憶部708と、標準モデル送信部709とを備える。

【0296】

参照モデル準備部702は、読み込み部711が読み込んだ、CD-ROMなどのストレージデバイスに書き込まれた話者・雑音・声の調子別の音声認識用参照モデルを参照モデル記憶部703へ送信し、参照モデル記憶部703は、送信された参照モデル721を記憶する。

【0297】

仕様情報受信部707は、端末装置712から仕様情報725を受信する。利用情報受信部704は、端末装置712から利用情報724である雑音下で発声した利用者の音声を受信する。参照モデル選択部705は、利用情報724である利用者の音声に音響的に近い話者・雑音・声調子の参照モデル723を、参照

モデル記憶部 703 が記憶している参照モデル 721 の中から選択する。

【0298】

標準モデル作成部 706 は、仕様情報受信部 707 が受信した仕様情報 725 に基づいて、参照モデル選択部 705 が選択した参照モデル 723 に対する確率または尤度を最大化または極大化するように標準モデル 722 を作成する処理部であり、第 2 の実施の形態における標準モデル作成部 206 と同一の機能を有する。標準モデル記憶部 708 は、仕様情報 725 に基づいた 1 もしくは複数の標準モデルを記憶する。標準モデル送信部 709 は、利用者の端末装置 712 から仕様情報と標準モデルの要求信号を受信すると、その仕様情報に適した標準モデルを端末装置 712 へ送信する。

【0299】

次に、以上のように構成されたサーバ 701 の動作について説明する。

図 26 は、サーバ 701 の動作手順を示すフローチャートである。図 27 は、サーバ 701 の動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【0300】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図 26 のステップ S700）。つまり、参照モデル準備部 702 は、読み込み部 711 が読み込んだ、CD-ROM などのストレージデバイスに書き込まれた話者・雑音・声の調子別の音声認識用参照モデルを参照モデル記憶部 703 へ送信し、参照モデル記憶部 703 は、送信された参照モデル 721 を記憶する。ここでは、参照モデル 721 は、話者・雑音・声の調子ごとに、音素ごとの HMM により構成される。また、各参照モデルは、図 27 の参照モデル 721 に示されるように、状態数 3 個、各状態は混合分布数が 128 個の混合ガウス分布により HMM の出力分布が構成される。特徴量として 25 次元 ($J=25$) のメルケプストラム係数が用いられる。

【0301】

次に、利用情報受信部 704 は、利用者 A の雑音下での音声を端末装置 712 から利用情報 724 として受信する（図 26 のステップ S701）。参照モデル

選択部 705 は、利用情報 724 である利用者 A の音声に音響的に近い参照モデル 723 を、参照モデル記憶部 703 が記憶している参照モデル 521 の中から選択する（図 26 のステップ S702）。具体的には、図 27 の「選択された参照モデル 723」に示されるように、ここでは、近い話者 100 人（ $N_g = 100$ ）の参照モデルが選択される

そして、仕様情報受信部 707 は、利用者 A の要求に基づき仕様情報 725 を端末装置 712 から受信する（図 26 のステップ S703）。ここでは、高い認識精度という仕様情報 725 を受信する。標準モデル作成部 706 は、仕様情報 725 に基づいて、参照モデル選択部 705 が選択した参照モデル 723 に対する確率または尤度を最大化または極大化するように標準モデル 722 を作成する（図 26 のステップ S704）。具体的には、標準モデル 722 は、図 27 の標準モデル 722 に示されるように、仕様情報 725 である高い認識精度という情報に基づいて、64 混合（ $M_f = 64$ ）で、3 状態の HMM より構成する。HMM は音素ごとに構成する。

【0302】

標準モデル 722 の作成方法は、第 2 の実施の形態と同様に行われる。

標準モデル記憶部 708 は、仕様情報 725 に基づいた 1 もしくは複数の標準モデル 722 を記憶する。ここでは、以前に作成した標準モデルである利用者 B の 16 混合の HMM がすでに記憶されており、新たに利用者 A の 64 混合の HMM が記憶される。

【0303】

利用者 A は、端末装置 712 からサーバ 701 の標準モデル送信部 709 へ、仕様情報である利用者 A と雑音の種類と標準モデルの要求信号とを送信する（図 26 のステップ S706）。標準モデル送信部 709 は、利用者 A が送信した仕様情報と標準モデルの要求信号とを受信すると、その端末装置 712 へ、仕様に適した標準モデルを端末装置 712 へ送信する（図 26 のステップ S707）。ここでは、先ほど作成した利用者 A の標準モデル 722 を端末装置 712 へ送信する。

【0304】

利用者Aは端末装置712において受信した標準モデル722を用いて音声認識を行う(図26のステップS708)。

以上説明したように、本発明の第7の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい精度の高い標準モデルが提供される。

【0305】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

また、標準モデル記憶部708は、複数の標準モデルを記憶することができるため、必要に応じてすぐに標準モデルが提供される。

【0306】

また、標準モデル送信部709により、標準モデルが端末装置712へ送信されるので、端末装置712とサーバ701が空間的に離れた場所に設置してある場合に、端末装置712は、容易にサーバ701が作成した標準モデルを利用することができる。

【0307】

なお、標準モデル722は、音素ごとにHMMを構成するに限らず、文脈依存のHMMで構成してもよい。

また、標準モデル722を構成するHMMは、状態ごとに異なる混合数の混合ガウス分布により構成してもよい。

【0308】

また、標準モデル722を用いて、サーバ701において音声認識を行い、認識結果を端末装置712へ送信してもよい。

また、参照モデル準備部702は、必要に応じて新たな参照モデルを作成して参照モデル記憶部703に追加・更新してもよいし、参照モデル記憶部703に格納されている不要な参照モデルを削除してもよい。

【0309】

また、参照モデル準備部702は、必要に応じて、通信路を介して新たな参照

モデルを参照モデル記憶部 703 に追加・更新してもよい。

また、標準モデルを作成したのちに、さらにデータにより学習してもよい。

【0310】

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などの HMM の構造や、状態数などを決定してもよい。

【0311】

(第 8 の実施の形態)

図 28 は、本発明の第 8 の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置が携帯電話 901 に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデルを作成する場合を例にして説明する。

【0312】

携帯電話 901 は、携帯情報端末であり、事象の集合と事象または事象間の遷移の出力確率で表現された隠れマルコフモデルによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、参照モデル受信部 909 と、参照モデル準備部 902 と、参照モデル記憶部 903 と、利用情報作成部 904 と、参照モデル選択部 905 と、類似度情報作成部 908 と、標準モデル作成部 906 と、仕様情報作成部 907 と、マイク 912 と、音声認識部 913 とを備える。

【0313】

利用情報作成部 904 は、利用情報 924 を携帯電話 901 の画面とキーを利用して作成する。

仕様情報作成部 907 は、携帯電話 901 の仕様に基づき仕様情報 925 を作成する。ここで、仕様情報とは、作成する標準モデルの仕様に関する情報であり、ここでは、携帯電話 901 が備える CPU の処理能力に関する情報である。

【0314】

類似度情報作成部 908 は、利用情報 924 と仕様情報 925 と参照モデル記憶部 903 が記憶した参照モデル 921 に基づいて、類似度情報 926 を作成して参照モデル準備部に送信する。

【0315】

参照モデル準備部 902 は、類似度情報 926 に基づいて、参照モデルを準備するか否かを決定する。参照モデル準備部 902 は、参照モデルを準備すると決定した場合に、利用情報 924 と仕様情報 925 を参照モデル受信部 909 に送信する。

【0316】

参照モデル受信部 909 は、利用情報 924 と仕様情報 925 に対応した参照モデルを、サーバ装置 910 から受信して参照モデル準備部 902 に送信する。

参照モデル準備部 902 は、参照モデル受信部 909 が送信した参照モデルを参照モデル記憶部 903 に記憶する。

【0317】

参照モデル選択部 905 は、利用情報 924 に対応した参照モデル 923 を、参照モデル記憶部 903 が記憶している参照モデル 921の中から選択する。

標準モデル作成部 906 は、仕様情報作成部 907 で作成された仕様情報 925 に基づいて、参照モデル選択部 905 が選択した参照モデル 923 に対する確率または尤度を最大化または極大化するように標準モデル 922 を作成する処理部であり、標準モデルの構造（ガウス分布の混合分布数など）を決定する標準モデル構造決定部 906a と、標準モデルを計算するための統計量の初期値を決定することで初期標準モデルを作成する初期標準モデル作成部 906b と、決定された初期標準モデルを記憶する統計量記憶部 906c と、統計量記憶部 906c に記憶された初期標準モデルに対して、第3近似部 906e による近似計算等を用いることにより、参照モデル選択部 905 が選択した参照モデル 923 に対する確率または尤度を最大化または極大化するような統計量を算出する（最終的な標準モデルを生成する）統計量推定部 906d とからなる。

【0318】

音声認識部 913 は、標準モデル作成部 906 で作成された標準モデル 922 を用いて、マイク 912 から入力された利用者の音声进行認識する。

次に、以上のように構成された携帯電話 901 の動作について説明する。

【0319】

図 29 は、携帯電話 901 の動作手順を示すフローチャートである。

いま、参照モデル記憶部 903 には、あらかじめ参照モデル 921 として子供用モデルが記憶されているとする。その参照モデル 921 は、音素ごとの HMM により構成される。参照モデル 921 の一例を図 30 に示す。ここでは、子供用参照モデルのイメージ図が示されている。これらの参照モデルは、状態数 3 個、各状態は分布数が 16 個の混合ガウス分布により HMM の出力分布が構成される。特徴量として、12 次元のメルケプストラム係数、12 次元のデルタメルケプストラム係数、デルタパワーの合計 25 次元 ($J = 25$) の特徴量が用いられる。

【0320】

まず、利用情報作成部 904 は、利用者の属するカテゴリである利用情報 924 を作成する (ステップ S900)。図 32 は、利用情報 924 の作成例を示す図である。図 32 (a) に携帯電話 901 の選択画面の一例を示す。ここでは、「4: 成人」のボタンを押すことにより、この携帯電話 901 が成人女性と成人男性に利用されることが選択されている。別の一例を図 32 (b) に示す。ここでは、「メニュー」ボタンを押しながら音声を入力している。その利用者の音声は、特徴量に変換されることで、利用情報 924 である「利用者の音声データ」が作成される。

【0321】

一方、仕様情報作成部 907 は、携帯電話 901 の仕様に基づき、仕様情報 925 を作成する (ステップ S901)。ここでは、携帯電話 901 のメモリ容量の大きさに基づいて「混合分布数 16」という仕様情報 925 を作成する。

【0322】

次に、類似度情報作成部 908 は、利用情報 924 と仕様情報 925 と参照モデル記憶部 903 が記憶した参照モデル 921 に基づいて、類似度情報 926 を作成して (ステップ S902)、類似度情報 926 を参照モデル準備部 902 に送信する。ここでは、参照モデル記憶部 903 に存在する参照モデル 921 は、混合分布数 3 の子供用モデル (図 30 を参照) のみであり、利用情報 924 である「成人」 (図 32 (a) に対応) と仕様情報 925 である「混合分布数 16」

に対応する参照モデルが参照モデル記憶部 903 に存在しないため、「類似した参照モデルが存在しない」という類似度情報 926 を作成して、類似度情報 926 を参照モデル準備部 902 に送信する。別の一例では、利用情報 924 は”利用者の音声データ”（図 32（b）に対応）であり、利用者の音声データを参照モデル記憶部 903 が記憶している子供用モデルに入力して類似度情報 926 を作成する。ここでは、子供用モデルに対する尤度が所定のしきい値以下であるため、「類似した参照モデルが存在しない」という類似度情報 926 を作成して参照モデル準備部 902 に送信する。

【0323】

続いて、参照モデル準備部 902 は、類似度情報 926 に基づいて、参照モデルを準備するか否かを決定する（ステップ S903）。ここでは、「類似した参照モデルが存在しない」ため、図 33（a）の携帯電話 901 の画面表示例に示すように利用者に参照モデルの準備を促す。ここで、利用者が「メモ」ボタンを押して参照モデルの準備を要求した場合に、参照モデル準備部 902 は、参照モデルを準備すると決定して、利用情報 924 と仕様情報 925 を参照モデル受信部 909 に送信する。別の一例では、「類似した参照モデルが存在しない」ため、参照モデル準備部 902 は、自動的に参照モデルを準備すると決定して、利用情報 924 と仕様情報 925 を参照モデル受信部 909 に送信する。この場合の携帯電話 901 の画面の一例を図 33（b）に示す。

【0324】

これに対して、参照モデル受信部 909 は、利用情報 924 と仕様情報 925 に対応した参照モデルをサーバ装置 910 から受信して参照モデル準備部 902 に送信する。ここでは、参照モデル受信部 909 は、利用情報 924 である「成人」（図 32（a）に対応）と仕様情報 925 である「混合分布数 16」に対応する参照モデルである、“混合分布数 16 の成人女性用モデル”と”混合分布数 16 の成人男性用モデル”の 2 個の参照モデルをサーバ装置 910 から受信する。

【0325】

そして、参照モデル準備部 902 は、参照モデル受信部 909 が送信した参照モデルを参照モデル記憶部 903 に記憶することによって参照モデルを準備する

(ステップS904)。図31にその参照モデルの一例を示す。ここでは、成人男性用、成人女性用、子供用の参照モデルのイメージ図が示されている。

【0326】

次に、参照モデル選択部905は、利用情報924である「成人」に対応した同じカテゴリに属する”混合分布数16の成人女性用モデル”と”混合分布数16の成人男性用モデル”の2個の参照モデルを参照モデル記憶部903が記憶している参照モデル921の中から選択する(ステップS905)。別の一例では、参照モデル選択部905は、利用情報924である”利用者の音声データ”と音響的に近い(尤度大きい)”混合分布数16の成人女性用モデル”と”混合分布数16の成人男性用モデル”の2個の参照モデルを参照モデル記憶部903が記憶している参照モデル921の中から選択する。

【0327】

続いて、標準モデル作成部906は、作成された仕様情報925に基づいて、参照モデル選択部905が選択した参照モデル923に対する確率または尤度を最大化または極大化するように標準モデル922を作成する(ステップS906)。

【0328】

最後に、音声認識部913は、標準モデル作成部906によって作成された標準モデル922に従って、マイク912から入力された利用者の音声を認識する(ステップS907)。

【0329】

次に、図29におけるステップS906(標準モデルの作成)の詳細な手順を説明する。手順の流れは、図4に示されたフローチャートと同様である。ただし、採用する標準モデルの構造や具体的な近似計算等が異なる。

【0330】

まず、標準モデル構造決定部906aは、標準モデルの構造を決定する(図4のステップS102)。ここでは、標準モデルの構造として、仕様情報925である「混合分布数16」に基づいて、音素ごとのHMMにより構成し、状態数を3とし、各状態における出力分布の混合分布数を16個($M_f=16$)と決定す

る。

【0331】

次に、初期標準モデル作成部 906b は、標準モデルを計算するための統計量の初期値を決定する（図4のステップ S102b）。ここでは、選択された参照モデル 923 である“混合分布数 16 の成人女性用モデル”を統計量の初期値として統計量記憶部 906c に記憶する。別の一例では、選択された参照モデル 923 である“混合分布数 16 の成人男性用モデル”を統計量の初期値として統計量記憶部 906c に記憶する。具体的には、初期標準モデル作成部 906b は、上記数 68 に示される出力分布を生成する。

【0332】

そして、統計量推定部 906d は、参照モデル選択部 905 が選択した 2 個の参照モデル 923 を用いて、統計量記憶部 906c に記憶された標準モデルの統計量を推定する（図4のステップ S102c）。つまり、2 個 ($N_g=2$) の参照モデル 923 における出力分布、即ち、上記数 74 に示される出力分布に対する標準モデルの確率（ここでは、上記数 80 に示される尤度 $\log P$ ）を極大化もしくは最大化するような標準モデルの統計量（上記数 71 に示される混合重み係数、上記数 72 に示される平均値、および、上記数 73 に示される分散値）を推定する。ただし、本実施の形態では、上記数 74 に示された出力分布における数 76 は、16（各参照モデルの混合分布数）である。

【0333】

具体的には、上記数 81、数 82 および数 83 に示される式に従って、それぞれ、標準モデルの混合重み係数、平均値および分散値を算出する。

【0334】

このとき、統計量推定部 906d の第 3 近似部 906e は、標準モデルの各ガウス分布はお互いに影響を与えないと仮定して、数 108 の近似式を用いる。また、繰り返し回数 R が 1 回目の場合には、数 109 に示される標準モデルのガウス分布の近傍の数 110 とは、数 109 が示す出力分布とのマハラノビス距離、カルバック・ライブラー (KL) 距離などの分布間距離が最も近いものと 2 番目に近いものの 2 個（近傍指示パラメータ $G=2$ ）の数 111 に示される参照モ

ル 9 2 3 のガウス分布が存在する空間であると近似する。一方、繰り返し回数 R が 2 回目以上の場合には、数 1 0 9 に示される標準モデルのガウス分布の近傍の数 1 1 0 とは、数 1 0 9 が示す出力分布とのマハラノビス距離、カルバック・ライブラー (KL) 距離などの分布間距離が最も近いもの 1 個 (近傍指示パラメータ $G=1$) の数 1 1 1 に示される参照モデル 9 2 3 のガウス分布が存在する空間であると近似する。

【0335】

以上の第 3 近似部 9 0 6 e による近似式を考慮してまとめると、統計量推定部 9 0 6 d での計算式は、次の通りになる。つまり、統計量推定部 9 0 6 d は、数 1 1 4、数 1 1 5 および数 1 1 6 に示される式に従って、それぞれ、混合重み係数、平均値および分散値を算出し、それらのパラメータによって特定される標準モデルを最終的な標準モデル 9 2 2 として生成する。ただし、第 3 の実施の形態における第 2 の方法である、混合重み係数の値をゼロにして、平均値をゼロ、分散値を 1 にする方法を用いる。また、繰り返し回数に対応して近傍指示パラメータ G の値は異なる。なお、近傍指示パラメータ G の値に依存して、上記の方法を、第 3 の実施の形態における第 1 から第 3 の方法のいずれかに決定してもよい。

【0336】

統計量推定部 9 0 6 d は、このように推定した標準モデルの統計量を統計量記憶部 9 0 6 c に記憶する。そして、このような統計量の推定と統計量記憶部 9 0 6 c への記憶を R (≥ 1) 回、繰り返す。その結果得られた統計量を最終的に生成する標準モデル 9 2 2 の統計量として出力する。

【0337】

図 3 4 に、第 3 近似部 9 0 6 e を用いて作成した標準モデル 9 2 2 を用いた認識実験の結果を示す。縦軸に成人 (男性と女性) の認識率 (%)、横軸に繰り返し回数 R を示す。繰り返し回数 $R=0$ とは、学習を行う前での初期標準モデル作成部 9 0 6 b が作成した初期モデルにより認識した結果である。また、繰り返し回数 $R=1$ のときは、近傍指示パラメータ $G=2$ とし、繰り返し回数 $R=2 \sim 5$ のときは、近傍指示パラメータ $G=1$ とした。

【0338】

グラフ「データ」は、数日間かけて音声データより学習した場合の結果を表しており、グラフ「女性」、グラフ「男性」は、それぞれ、初期モデルを成人女性、成人男性としたときの結果を表している。参照モデルによる本発明による学習時間は数十秒のオーダーであった。実験結果より、短時間に高い精度の標準モデルが作成できていることがわかる。

【0339】

ここで、参考のために、図35に、第3の実施の形態における第2近似部306eにより作成された標準モデルによる認識率を示す。本実施の形態における第3近似部906eと異なるのは、繰り返し回数Rによらず近傍指示パラメータ $G=1$ であるということである。実験結果より、初期モデルとして成人女性を選択すると良好な結果が得られることがわかる。また、初期モデルとして成人男性を選択すると、精度が少し劣化していることがわかる。図34の結果とあわせると、第3近似部906eによる標準モデルは初期モデルに依存せずに高い精度の標準モデルが作成できていることがわかる。

【0340】

以上説明したように、本発明の第8の実施の形態によれば、類似度情報に基づいて参照モデルを準備するため、利用情報および仕様情報にふさわしい参照モデルを必要なタイミングで準備することができる。また、近傍指示パラメータ G を繰り返し回数Rによって変化させることで、初期モデルにかかわらず精度の高い標準モデルを提供することができる。

【0341】

なお、統計量推定部906dによる処理の繰り返し回数は、上記数80に示された尤度の大きさがある一定のしきい値以上になるまでの回数としてもよい。

【0342】

また、標準モデル922は、音素ごとにHMMを構成するに限らず、文脈依存のHMMで構成してもよい。

また、標準モデル作成部906は、一部の音素の、一部の状態における事象の出力確率に対してモデル作成を行ってもよい。

【0343】

また、標準モデル 922 を構成する HMM は、音素ごとに異なる状態数により構成してもよいし、状態ごとに異なる分布数の混合ガウス分布により構成してもよい。

【0344】

また、標準モデルを作成したのちに、さらに音声データにより学習してもよい。

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などの HMM の構造や、状態数などを決定してもよい。

【0345】

(第 9 の実施の形態)

図 36 は、本発明の第 9 の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置が PDA (Personal Digital Assistant) 1001 に組み込まれた例が示されている。以下、本実施の形態では音声認識用の標準モデルを作成する場合を例にして説明する。

【0346】

PDA 1001 は、携帯情報端末であり、事象の集合と事象または事象間の遷移の出力確率で表現された隠れマルコフモデルによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、参照モデル記憶部 1003 と、標準モデル作成部 1006 と、アプリ・仕様情報対応データベース 1014 と、マイク 1012 と、音声認識部 1013 とを備える。標準モデル作成部 1006 は、標準モデル構造決定部 1006a と、初期標準モデル作成部 1006b と、統計量記憶部 306c と、統計量推定部 306d とを備える。

【0347】

標準モデル作成部 1006 は、送信されたアプリ起動情報 1027 (ここでは、起動したアプリケーションの ID 番号) に基づいて、アプリ・仕様情報対応データベース 1014 を用いて、仕様情報 1025 を取得する。図 37 は、仕様情報対応データベース 1014 のデータ例を示す。仕様情報対応データベース 1014 には、アプリケーション (ID 番号および名前) に対応する仕様情報 (こ

では、混合分布数)が登録されている。

【0348】

標準モデル作成部1006は、取得した仕様情報1025に基づいて、参照モデル記憶部1003が記憶した1個の参照モデル1023に対する確率または尤度を最大化または極大化するように標準モデル1022を作成する処理部であり、第3の実施の形態における第2近似部306eの機能を有する。

【0349】

音声認識部1013は、標準モデル作成部1006で作成された標準モデル1022を用いて、マイク1012から入力された利用者の音声を認識する。

次に、以上のように構成されたPDA1001の動作について説明する。

【0350】

図38は、PDA1001の動作手順を示すフローチャートである。

ここで、参照モデル記憶部1003には、あらかじめ多くの混合分布数をもつ利用者用モデルが参照モデル1021として1個、記憶されているとする。参照モデル1021は、音素ごとのHMMにより構成される。参照モデル1021の一例を図39に示す。この参照モデルは、状態数3個、各状態は分布数が300個の混合ガウス分布によりHMMの出力分布が構成される。特徴量として、12次元のメルケプストラム係数、12次元のデルタメルケプストラム係数、デルタパワーの合計25次元($J=25$)の特徴量が用いられる。

【0351】

まず、利用者は、例えば「株取引」というアプリケーションを起動する(ステップS1000)。

これに対して、標準モデル作成部1006は、アプリ起動情報として起動されたアプリケーションのID「3」を受信する(ステップS1001)。そして、アプリ・仕様情報対応データベース1014を用いてID「3」に対応する仕様情報1025である「混合分布数126」に基づいて、標準モデル1022を作成する(ステップS1002)。具体的には、標準モデル1022として、混合分布数126($M_f=126$)で、3状態の文脈依存型のHMMにより構成する。

。

【0352】

次に、標準モデル作成部1006は、仕様情報1025を受信して（ステップS1001）、仕様情報1025に基づいて標準モデルを作成する（ステップS1002）。

【0353】

最後に、音声認識部1013は、標準モデル作成部1006によって作成された標準モデル1022に従って、マイク1012から入力された利用者の音声进行認識する（ステップS1003）。

【0354】

次に、図38におけるステップS1002（標準モデルの作成）の詳細な手順を説明する。手順の流れは、図4に示されたフローチャートと同様である。ただし、採用する標準モデルの構造や具体的な近似計算等が異なる。

【0355】

まず、標準モデル構造決定部1006aは、アプリ起動情報1027としてアプリケーションID「3」を受信した後に、アプリ・仕様情報対応データベース1014を用いてID「3」に対応した仕様情報1025（「混合分布数126」）を参照することにより、標準モデルの構造を混合分布数126（ $M_f=126$ ）で、3状態の文脈依存型のHMMと決定する（図4のステップS102a）。

【0356】

そして、初期標準モデル作成部1006bは、標準モデル構造決定部1006aが決定した標準モデルの構造に基づいて、標準モデルを計算するための統計量の初期値を決定する（図4のステップS102b）ここでは、k-means法とマハラノビス汎距離を用いた方法により、後述するクラスタリングを行ったものを統計量の初期値として統計量記憶部306cに記憶する。

【0357】

そして、統計量推定部306dは、参照モデル記憶部1003に格納された参照モデル1021を用いて、統計量記憶部306cに記憶された標準モデルの統計量を推定する（図4のステップS102c）。なお、この統計量推定部306

dによる推定処理は、第3の実施の形態と同様である。

【0358】

次に、初期標準モデル作成部1006bによる初期値の決定方法、つまり、k-means法とマハラノビス汎距離を用いた方法によるクラスタリングについて説明する。図40にクラスタリングのフローチャートを示す。また、図41～図44にクラスタリングのイメージ図を示す。

【0359】

まず、図40のステップS1004において、標準モデルの混合分布数である126個の代表点を準備する(図41)。ここでは、参照モデルの300個の出力分布の中から126個の出力分布を選択して、選択された分布の平均値を代表点とする。

【0360】

次に、図40のステップS1005において、各代表点にマハラノビス汎距離が近い参照モデルの出力ベクトルを決定する(図42)。そして、図40のステップS1006において、ステップS1005で決定した近い分布を1つのガウス分布で表現して平均値を新しい代表点とする(図43)。

【0361】

続いて、図40のステップS1007において、クラスタリング操作を停止するかどうかを決定する。ここでは、各代表点と参照ベクトルの分布とのマハラノビス汎距離の変化率(1回前の代表点との距離との差分)がしきい値以下になった場合に停止とする。停止条件を満たさない場合、図40のステップS1005に戻り、近い分布を決定して同様の操作を繰り返す。

【0362】

一方、停止条件を満たす場合には、図40のステップS1008に進み、統計量の初期値を決定して統計量記憶部906cに記憶する。このようにして、クラスタリングによる初期値の決定が行われる。

【0363】

以上説明したように、本発明の第9の実施の形態によれば、アプリケーションに連動して自動的に仕様情報にふさわしい標準モデルを獲得することができる。

なお、標準モデル1022は、音素ごとにHMMを構成してもよい。

【0364】

また、標準モデル作成部1006は、一部の音素の、一部の状態における事象の出力確率に対してモデル作成を行ってもよい。

また、標準モデル1022を構成するHMMは、音素ごとに異なる状態数により構成してもよいし、状態ごとに異なる分布数の混合ガウス分布により構成してもよい。

【0365】

また、標準モデルを作成したのちに、さらに音声データにより学習してもよい。

また、標準モデル構造決定部は、モノフォン、トライフォン、状態共有型などのHMMの構造や、状態数などを決定してもよい。

【0366】

(第10の実施の形態)

図45は、本発明の第10の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ801に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデル（適応モデル）を作成する場合を例にして説明する。

【0367】

サーバ801は、通信システムにおけるコンピュータ装置等であり、事象の集合と事象または事象間の遷移の出力確率とによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、読み込み部711と、参照モデル準備部702と、参照モデル記憶部703と、利用情報受信部704と、参照モデル選択部705と、標準モデル作成部706と、仕様情報受信部707と、標準モデル記憶部708と、標準モデル送信部709と、参照モデル受信部810とを備える。

【0368】

参照モデル準備部702は、読み込み部711が読み込んだ、CD-ROMな

どのストレージデバイスに書き込まれた話者・雑音・声の調子別の音声認識用参照モデルを参照モデル記憶部 703 へ送信する。参照モデル記憶部 703 は、送信された参照モデル 721 を記憶する。また、参照モデル準備部 702 は、端末装置 712 からの送信に対して参照モデル受信部 810 が受信した音声認識用参照モデルを参照モデル記憶部 703 へ送信する。参照モデル記憶部 703 は、送信された参照モデル 721 を記憶する。

【0369】

仕様情報受信部 707 は、端末装置 712 から仕様情報 725 を受信する。利用情報受信部 704 は、端末装置 712 から利用情報 724 である雑音下で発声した利用者の音声を受信する。参照モデル選択部 705 は、利用情報受信部 704 が受信した利用情報 724 である利用者の音声に音響的に近い話者・雑音・声調子の参照モデル 723 を、参照モデル記憶部 703 が記憶している参照モデル 721 から選択する。

【0370】

標準モデル作成部 706 は、仕様情報 725 に基づいて、参照モデル選択部 705 が選択した参照モデル 723 に対する確率または尤度を最大化または極大化するように標準モデル 722 を作成する処理部であり、第 2 の実施の形態における標準モデル作成部 206 と同一の機能を有する。標準モデル記憶部 708 は、仕様情報 725 に基づいた 1 もしくは複数の標準モデルを記憶する。標準モデル送信部 709 は、利用者の端末装置 712 から、仕様情報と標準モデルの要求信号とを受信すると、その端末装置 712 へ、仕様に適した標準モデルを送信する。

【0371】

次に、以上のように構成されたサーバ 801 の動作について説明する。

図 46 は、サーバ 801 の動作手順を示すフローチャートである。なお、このサーバ 801 の動作手順を説明するための参照モデルおよび標準モデルの一例は、第 7 に実施の形態における図 27 と同様である。

【0372】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図

46のステップS800、S801)。つまり、参照モデル準備部702は、読み込み部711が読み込んだ、CD-ROMなどのストレージデバイスに書き込まれた話者・雑音・声の調子別の音声認識用参照モデルを参照モデル記憶部703へ送信し、参照モデル記憶部703は、送信された参照モデル721を記憶する(図46のステップS800)。ここでは、参照モデル721は、話者・雑音・声の調子ごとに、音素ごとのHMMにより構成される。また、参照モデル準備部702は、端末装置712が送信して参照モデル受信部810が受信した、利用者と端末装置712に適した音声認識用参照モデルを参照モデル記憶部703へ送信し、参照モデル記憶部703は、送信された参照モデル721を記憶する(図46のステップS801)。ここでは、各参照モデルは、図27の参照モデル721に示されるように、状態数3個、各状態は混合分布数が128個の混合ガウス分布によりHMMの出力分布が構成される。特徴量として25次元($J=25$)のメルケプストラム係数が用いられる。

【0373】

以下、これらの参照モデル721を用いた標準モデル722の作成および端末装置712への送信(図46のステップS802~S809)は、第7の実施の形態における手順(図26のステップS701~S708)と同様である。

【0374】

図47は、本実施の形態における標準モデル作成装置を具体的に適用したシステム例を示す図である。ここには、インターネットや無線通信等を介して通信し合うサーバ701と端末装置712(携帯電話機712a、カーナビゲーション装置712b)とが示されている。

【0375】

たとえば、携帯電話機712aは、利用者の音声を利用情報とし、携帯電話機での利用である旨(CPUの処理能力が低いこと)を仕様情報とし、予め記憶しているサンプルモデルを参照モデルとし、それら利用情報、仕様情報および参照モデルをサーバ701に送信することで、標準モデルの作成を要求する。その要求に対してサーバ701で標準モデルが作成されると、携帯電話機712aは、その標準モデルをダウンロードし、その標準モデルを用いて利用者の音声を認識

する。例えば、利用者の音声で、内部に保持するアドレス帳の名前と一致した場合には、その名前に対応する電話番号に自動発呼する。

【0376】

また、カーナビゲーション装置 712b は、利用者の音声を利用情報とし、カーナビゲーション装置での利用である旨（CPU の処理能力が通常であること）を仕様情報とし、予め記憶しているサンプルモデルを参照モデルとし、それら利用情報、仕様情報および参照モデルをサーバ 701 に送信することで、標準モデルの作成を要求する。その要求に対してサーバ 701 で標準モデルが作成されると、カーナビゲーション装置 712b は、その標準モデルをダウンロードし、その標準モデルを用いて利用者の音声を認識する。例えば、利用者の音声で、内部に保持する地名と一致した場合には、その地名を目標点とする現地点からの道順を示す地図を画面に自動表示する。

【0377】

このようにして、携帯電話機 712a およびカーナビゲーション装置 712b は、自装置に適した標準モデルの作成をサーバ 701 に依頼することで、標準モデルの作成に必要な回路や処理プログラムを自装置内に実装する必要がなくなるとともに、様々な認識対象の標準モデルを必要なタイミングで獲得することができる。

【0378】

以上説明したように、本発明の第 10 の実施の形態によれば、参照モデル受信部 810 が受信した参照モデルを利用して標準モデルを作成できるため、精度の高い標準モデルが提供される。

【0379】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

なお、参照モデル受信部 810 は、端末装置 712 とは異なる他の端末装置から参照モデルを受信してもよい。

【0380】

また、図 47 に示された応用例は、本実施の形態に限られるものではなく、他

の実施の形態にも適用することができる。つまり、第1～第9の実施の形態で作成された標準モデルを各種記録媒体や通信を介して様々な電子機器に配信することで、それらの電子機器において、制度の高い音声認識、画像認識、意図理解等を行うことが可能となる。さらに、上記実施の形態における標準モデル作成装置を各種電子機器に内蔵させることで、音声認識、画像認識、意図理解等の認識・認証機能を備えるスタンドアローンの電子機器を実現することもできる。

【0381】

以上、本発明に係る標準モデル作成装置について、実施の形態に基づいて説明したが、本発明は、これらの実施の形態に限定されるものではない。

たとえば、第1～第10の実施の形態における標準モデルの統計量の近似計算については、各実施の形態における近似計算だけに限られず、第1～第4の実施の形態における合計4種類の近似計算の少なくとも1つを用いてもよい。つまり、4種類の近似計算のいずれであってもよいし、2以上の種類の近似計算の組み合わせであってもよい。

【0382】

また、初期標準モデル作成部による初期標準モデルの作成においては、図48に示されるようなクラスID・初期標準モデル・参照モデル対応表を予め準備しておき、この表に従って、初期標準モデルを決定してもよい。以下、このようなクラスID・初期標準モデル・参照モデル対応表を用いた初期標準モデルの決定方法について説明する。なお、クラスIDとは、標準モデルを用いた認識対象の種別を識別するIDであり、標準モデルの種類に対応する。

【0383】

図48に示されたクラスID・初期標準モデル・参照モデル対応表は、一定の共通する性質を有する複数の参照モデルに対して、それらを識別する1つのクラスIDを対応づけるとともに、それら参照モデルと共通する性質を持つ予め作成された初期標準モデルを対応づけた表である。この表では、参照モデル8AA～8AZに対して、クラスIDおよび初期標準モデル8Aが対応づけられ、参照モデル64ZA～ZZに対して、クラスIDおよび初期標準モデル64Zが対応づけられている。標準モデル作成部は、使用する参照モデルの性質と共通する初期

標準モデルを使用することによって、精度の高い標準モデルを生成することができる。

【0384】

ここで、クラスID、初期標準モデルおよび参照モデルの添え字記号8A、8AAにおける最初の記号「8」等は、混合分布数を意味し、2番目の記号「A」等は大分類、例えば、騒音下における音声認識の場合であれば、騒音環境の種類（家庭内騒音下をA、電車内騒音下をBなど）を意味し、3番目の記号「A」等は小分類、例えば、音声認識の対象となる人の属性（低学年の小学生をA、高学年の小学生をBなど）を意味する。したがって、図48のクラスID・初期標準モデル・参照モデル対応表における参照モデル8AA～AZは、図49に示されるような混合分布数8のモデルであり、参照モデル64ZA～ZZは、図50に示されるような混合分布数64のモデルであり、初期標準モデル8A～64Zは、図51に示されるような混合分布数8～16のモデルである。

【0385】

次に、このようなクラスID・初期標準モデル・参照モデル対応表の作成方法を説明する。図52は、その手順を示すフローチャートであり、図53～図56は、各ステップでの具体例を示す図である。ここでは、騒音環境下での音声認識を例とし、表だけでなく、クラスID、初期標準モデルおよび参照モデルも含めて新規に作成する場合の手順を説明する。

【0386】

まず、音声データを音響的に近いグループに分類する（図52のステップS1100）。たとえば、図53に示されるように、音声データを利用情報である雑音環境で分類する。環境A（家庭内騒音下での音声データ）には、家庭内騒音下で収録した小学生低学年の音声、小学生高学年の音声、成人女性の音声などが含まれ、環境B（電車内での音声データ）には、電車内で収録した小学生低学年の音声、小学生高学年の音声、成人女性の音声などが含まれるように分類する。なお、利用情報である話者の性別、年齢層、笑い声・怒った声などの声の性質、読み上げ調・会話調などの声の調子、英語・中国語などの言語などで分類してもよい。

【0387】

次に、仕様情報等に基づいて、準備する参照モデルの1以上のモデル構造を決定する（図52のステップS1101）。たとえば、8混合、16混合、32混合および64混合を対象とすることを決定する。なお、モデル構造の決定においては、混合分布数を決定するに限らず、HMMの状態数、モノフォン・トライフォンなどのHMMの種類などを決定してもよい。

【0388】

続いて、初期標準モデルを作成する（図52のステップS1102）。つまり、上記音声データの分類（ステップS1100）において決定した分類（環境A、環境B、…）ごとに、ステップS1101において決定したモデル構造ごとの初期標準モデルを作成する。例えば、図54に示されるように、初期標準モデル8Aであれば、8混合の初期標準モデルを、家庭内騒音下（環境A）における音声データ（低学年の小学生、高学年の小学生、成人男、成人女等の音声データ）を用いて、バウム・ウェルチアルゴリズムなどにより学習して作成する。

【0389】

次に、参照モデルを作成する（図52のステップS1103）。つまり、上記ステップS1102において作成した初期標準モデルを用いて参照モデルを作成する。具体的には、参照モデルを学習する音声データの雑音環境と同じ雑音環境で学習した、同じ混合分布数をもつ初期標準モデルを用いて参照モデルを学習する。例えば、図55に示されるように、参照モデル8AAは、混合分布数8の家庭内騒音下での小学生低学年の音声データで学習するモデルであり、学習を行う際の初期値として、同じ環境である家庭内騒音下での音声データ（小学生低学年、小学生高学年、成人女性、成人男性の音声を含む）で学習した初期標準モデルを用いる。学習方法として、バウム・ウェルチアルゴリズムを用いる。

【0390】

最後に、クラスIDを付与する（図52のステップS1104）。たとえば、騒音環境下ごとに1つのクラスIDを付与することによって、図56に示されるクラスID・初期標準モデル・参照モデル対応表、つまり、“クラスID付き初期標準モデル”および“クラスID付き参照モデル”が作成される。

【0391】

なお、このようなクラスID・初期標準モデル・参照モデル対応表は、完成された表として予め端末（標準モデル作成装置）が保持している必要はない。端末（標準モデル作成装置）は、図57に示されるように、他の装置（サーバ）と通信することによって表を完成させてもよい。つまり、標準モデル作成装置（端末）は、通信網などを介して、“クラスID付き初期標準モデル”、“クラスID付き参照モデル”を取得することが可能である。もっとも、端末は必ずしも“クラスID付き初期標準モデル”、“クラスID付き参照モデル”を取得する必要はなく事前に記憶させて出荷してもよい。

【0392】

図57に示されるように、端末は、以下のような方法によって、“クラスID付き初期標準モデル”、“クラスID付き参照モデル”を取得することができる。第1の方法として、端末は、“クラスID付き初期標準モデル”（例えば規格化コンソーシアムなどで事前に定義されたクラスIDのつけ方に遵守したもの）を記憶しているケースである。このとき、端末は、1以上のサーバから“クラスID付き参照モデル”（例えば規格化コンソーシアムなどで事前に定義されたクラスIDのつけ方に遵守したもの）をダウンロードする。なお、端末に、“クラスID付き参照モデル”を出荷時に記憶させておいてもよい。

【0393】

また、第2の方法として、端末は、“クラスID付き初期標準モデル”を記憶していないケースである。このとき、端末は、サーバ（図57のサーバ1）から“クラスID付き初期標準モデル”をダウンロードする。次に、端末は、1以上のサーバ（図57のサーバ2）から“クラスID付き参照モデル”をダウンロードする。必要に応じて逐次的にクラスIDの定義の追加、変更が可能である。また、端末のメモリの節約にもなる。

【0394】

さらに、第3の方法として、端末は、クラスIDと初期標準モデル・参照モデルの対応関係を明記した“クラスID・初期標準モデル・参照モデル対応表”を記憶しているケースである。このとき、端末は、“対応表”記憶していないサーバ（

図 57 のサーバ 3) に”対応表”をアップロードする。サーバは、送信された”対応表”に基づき”クラス ID 付き参照モデル”を準備する。端末は、準備された”クラス ID 付き参照モデル”をダウンロードする。

【0395】

次に、このようなクラス ID・初期標準モデル・参照モデル対応表を用いた初期標準モデル作成部による初期標準モデルの決定方法について説明する。図 58 は、その手順を示すフローチャートである。図 59 および図 60 は、各ステップでの具体例を示す図である。

【0396】

まず、標準モデルの作成に用いる参照モデルからクラス ID を抽出する（図 58 のステップ S1105）。たとえば、図 59 に示されるテーブルに従って、選択された参照モデルから、対応するクラス ID を抽出する。ここでは、抽出したクラス ID として、8A が 1 個、16A が 3 個、16B が 1 個、64B が 1 個とする。

【0397】

次に、抽出したクラス ID を用いて標準モデル作成に用いる初期標準モデルを決定する（図 58 のステップ S1106）。具体的には、以下の手順に従って初期標準モデルを決定する。

(1) 作成する標準モデルの混合分布数（16 混合）と同じクラス ID（16*）をもつ参照モデルから抽出したクラス ID（16A、16B）に着目し、その中から一番多く抽出されたクラス ID に対応する初期標準モデルを最終的な初期標準モデルと決定する。たとえば、標準モデルの構造が 16 混合の場合には、16 混合に関するクラス ID として、16A が 3 個、16B が 1 個抽出されているので、クラス ID が 16B の初期標準モデルを採用する。

(2) 作成する標準モデルの混合分布数（8 混合）と同じクラス ID（8*）をもつ参照モデルから抽出したクラス ID（8A）に着目し、同じクラス ID をもつ初期標準モデルを最終的な初期標準モデルと決定する。たとえば、標準モデルの構造が 8 混合の場合には、8 混合に関するクラス ID として、8A が 1 個抽出されているので、クラス ID が 8A の初期標準モデルを採用する。

(3) 作成する標準モデルの混合分布数(32混合)と同じクラスID(32*)をもつ参照モデルから抽出したクラスIDに着目し、存在しない場合、仕様情報に着目してその中から一番多く抽出されたクラスID(*A)をもつ初期標準モデル(8A、16A)を用いてクラスタリングにより32混合にして最終的な初期標準モデルとする(図40を参照)。たとえば、標準モデルの構造が32混合の場合には、32混合に関するクラスIDが抽出されていないので、一番多く抽出されたクラスID(16A)を用いてクラスタリングにより32混合にして初期標準モデルとする。

【0398】

なお、はじめに作成する標準モデルの仕様情報(混合分布数など)に着目せず、利用情報(雑音の種類など)に着目して初期値を決定してもよい。

図60に、第3近似部を用いて作成した混合分布数が64の標準モデルを用いた認識実験の結果を示す。縦軸に成人(男性と女性)の認識率(%)、横軸に繰り返し回数Rを示す。繰り返し回数 $R=0$ とは、学習を行う前での初期標準モデル作成部が作成した初期モデルにより認識した結果である。また、繰り返し回数 $R=1\sim5$ において、近傍指示パラメータ $G=1$ とした。

【0399】

グラフ「データ」は、数日間かけて音声データより学習した場合の結果を表しており、グラフ「女性」、グラフ「男性」は、それぞれ、初期モデルを成人女性、成人男性としたときの結果を表している。参照モデルによる本発明による学習時間は数分のオーダーであった。この実験結果より、成人女性の参照モデルを初期標準モデルと決定した場合には、音声データで学習した結果よりも高い精度の標準モデルが作成できていることが分かる。

【0400】

このことは、音声データを分割し、分割した音声データをそれぞれの参照モデルとして厳密に学習したのちに統合したほうが、音声データによる学習の課題である局所解に陥るという問題を解決できる可能性を示している(音声データによる学習との認識精度での比較)。

【0401】

また、音声データの収録が困難な子供の音声データに対しては、データ数に適切である混合分布数の少ない参照モデルで厳密に学習して、多くの音声データの収録が可能な成人の音声データに対しては、混合分布数の多い参照モデルで厳密に学習して、そのあとで本発明により統合して標準モデルを作成すれば、極めて精度の高い標準モデルが作成できることが期待できる。

【0402】

なお、標準モデルの混合分布数が16の場合における認識実験(図35)では、本発明による方法は、音声データで学習した標準モデルの認識率を超えていない。このことは、音声データを16混合の参照モデルの形にしたときに音声データの情報が欠如したためだと考えられる。参照モデルを64混合で作成して音声データの特徴を十分保持しておけばより高い精度の標準モデルが作成できる。このことより、第9の実施の形態では、参照モデルの混合分布数を300と大きめに設定している。

【0403】

また、図35および図60に示される認識実験より、初期標準モデルが認識精度に与える影響が示されており、初期標準モデルの決定方法の重要性を物語っている(図60において、成人女性の参照モデルを初期標準モデルとして利用した場合、成人男性の参照モデルを利用する場合より高い精度の標準モデルが作成できることが示されている)。

【0404】

以上のように、クラスID・初期標準モデル・参照モデル対応表に従って、参照モデルと共通する性質の初期標準モデルを用いることで、精度の高い標準モデルを作成することができる。

【0405】

なお、このようなクラスID・初期標準モデル・参照モデル対応表を用いた初期標準モデルの決定は、上記実施の形態1~10のいずれにおいても採用することができる。

【0406】

【発明の効果】

以上の説明から明らかなように、本発明に係る標準モデル作成装置によれば、1以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように標準モデルの統計量が計算され、標準モデルが作成されるので、学習データや教師データを必要とすることなく簡易に標準モデルが作成されるとともに、既に作成された複数の参照モデルを総合的に勘案した高精度な標準モデルが作成される。

【0407】

また、標準モデル作成装置の外部から新たな参照モデルを取り込み、取り込んだ参照モデルに基づいた標準モデルが作成されるので、様々な認識対象に対応した汎用性の高い標準モデル作成装置が実現される。

【0408】

また、利用者の特徴、利用者の年齢、性別、利用環境などの利用情報に基づいて選択された参照モデルを基準として、標準モデルが作成されるので、認識対象により特化した精度の高い標準モデルが作成される。

【0409】

また、通信路を介して送信されてきた参照モデルに基づいて標準モデルが作成されるので、遠隔制御による標準モデルの生成が可能になるとともに、通信システムを基盤とする認識システムの構築が実現される。

【0410】

また、標準モデルを使用する装置のCPUパワー、記憶容量、要求される認識精度、要求される認識処理時間などの仕様情報に基づいて標準モデルが作成されるので、特定の仕様条件を満たす標準モデルの生成が可能となり、計算エンジン等の認識処理に必要なリソース環境に適した標準モデルの生成が実現される。

【0411】

また、通信路を介して送信されてきた仕様情報に基づいて標準モデルが作成されるので、遠隔制御による標準モデルの生成が可能になるとともに、通信システムを基盤とする認識システムの構築が実現される。

【0412】

また、仕様情報に基づいて、標準モデルに含まれるガウス分布の混合分布数が

動的に決定されるので、認識処理が実行される環境や要求仕様等に応じて標準モデルの構造を制御することが可能となる。たとえば、標準モデルを使用する認識装置のCPUパワーが小さい場合、記憶容量が小さい場合、要求される認識処理時間が短い場合などは標準モデルの混合数を少なく設定して仕様に合わせることができ、一方、要求される認識精度が高い場合などは混合数を多く設定して認識精度を高くすることができる。

【0413】

また、仕様情報が各アプリケーションに対応づけられているので、アプリケーションごとに最適な標準モデルが作成され、標準モデルが使用される認識システム等における認識精度が向上される。

【0414】

また、混合分布数が異なる参照モデルに基づいて標準モデルが作成されるので、予め準備された多種多様な構造の参照モデルに基づく標準モデルの作成が可能となり、より認識対象に適した精度の高い標準モデルの作成が実現される。

【0415】

また、作成された標準モデルは内部の記憶手段に保持されるので、作成された標準モデルを一時的にバッファリングしておき、送信要求に対してすぐに出力したり、他の装置に提供するデータサーバとしての役割を果たしたりすることが可能となる。

【0416】

また、作成された標準モデルは空間的に離れた場所に置かれた外部装置に送信されるので、本標準モデル作成装置を標準モデル作成エンジンとして独立させたり、標準モデル作成装置を通信システムにおけるサーバとして機能させたりすることが可能になる。

【0417】

また、準備される参照モデルの追加、更新等が行われるので、様々な認識対象用のモデルを参照モデルとして追加したり、より精度の高い参照モデルに置き換えたりすることが可能となり、更新した参照モデルによる標準モデルの再生成や、生成された標準モデルを参照モデルとして再び標準モデルを作成するというフ

ィードバックによる学習等が可能となる。

【0418】

また、標準モデルが使用される認識の対象の種類ごとに付与されたクラスIDに基づいて初期標準モデルが決定され、その初期標準モデルに基づいて標準モデルが作成されるので、最終的に必要とされる標準モデルと共通の性質をもつ初期標準モデルが使用され、短時間で精度の高い標準モデルが作成される。

【0419】

以上のように、本発明により、隠れマルコフモデル、ベイズ理論、線形判別分析などの確率モデルによる音声認識、文字認識、画像認識などのパターン認識、ベイジアンネットなどの確率モデルによる意図理解（意図の認識）、確率モデルによるデータマイニング（データ特性の認識）、確率モデルによる人物検出、指紋認証、顔認証、虹彩認証（対象を認識して特定の対象かどうかを判断する）、株価予測、天気予測などの予測（状況を認識して判断する）などに用いる高精度な標準モデルが提供され、その実用的価値は極めて高い。

【図面の簡単な説明】

【図1】

本発明の第1の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図2】

同サーバの動作手順を示すフローチャートである。

【図3】

図1における参照モデル記憶部に記憶されている参照モデルの例を示す図である。

【図4】

図2におけるステップS101（標準モデルの作成）の詳細な手順を示すフローチャートである。

【図5】

図1における第1近似部104eによる近似計算を説明する図である。

【図6】

本発明の第2の実施の形態における標準モデル作成装置に係るSTBの全体構成を示すブロック図である。

【図7】

同STBの動作手順を示すフローチャートである。

【図8】

図7における参照モデル記憶部に記憶されている参照モデルの例を示す図である。

【図9】

図7における第2近似部による近似計算を説明する図である。

【図10】

本発明の第3の実施の形態における標準モデル作成装置に係るPDAの全体構成を示すブロック図である。

【図11】

同PDAの動作手順を示すフローチャートである。

【図12】

図10における参照モデル記憶部に記憶されている参照モデルの例を示す図である。

【図13】

同PDAの選択画面の一例を示す。

【図14】

図10における統計量推定部による統計量の推定手順を示す概念図である。

【図15】

図10における第3近似部による近似計算を説明する図である。

【図16】

本発明の第4の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図17】

同サーバの動作手順を示すフローチャートである。

【図18】

同サーバの動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【図 19】

本発明の第 5 の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図 20】

同サーバの動作手順を示すフローチャートである。

【図 21】

同サーバの動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【図 22】

本発明の第 6 の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図 23】

同サーバの動作手順を示すフローチャートである。

【図 24】

同サーバの動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【図 25】

本発明の第 7 の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図 26】

同サーバの動作手順を示すフローチャートである。

【図 27】

同サーバの動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【図 28】

本発明の第 8 の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。

【図 29】

携帯電話 901 の動作手順を示すフローチャートである。

【図 30】

参照モデル記憶部に格納されている参照モデルの一例を示す図である。

【図 31】

新たに参照モデル記憶部に格納された参照モデルの一例を示す図である。

【図 32】

利用情報を作成するときの画面表示例を示す図である。

【図 33】

参照モデルを準備するときの画面表示例を示す図である。

【図 34】

第 3 近似部を用いて作成した標準モデルを用いた認識実験の結果を示すグラフである。

【図 35】

第 3 の実施の形態における第 2 近似部により作成された標準モデルによる認識実験の結果を示すグラフである。

【図 36】

本発明の第 9 の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。

【図 37】

アプリ・仕様情報対応データベースのデータ例を示す図である。

【図 38】

PDA1001 の動作手順を示すフローチャートである。

【図 39】

参照モデル記憶部に格納されている参照モデルの一例を示す図である。

【図 40】

初期標準モデル作成部によるクラスタリングによる初期値の決定方法を示すフローチャートである。

【図 4 1】

図 4 0 におけるステップ S 1 0 0 4 の具体例を示す図である。

【図 4 2】

図 4 0 におけるステップ S 1 0 0 5 の具体例を示す図である。

【図 4 3】

図 4 0 におけるステップ S 1 0 0 6 の具体例を示す図である。

【図 4 4】

図 4 0 におけるステップ S 1 0 0 8 の具体例を示す図である。

【図 4 5】

本発明の第 1 0 の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図 4 6】

同サーバの動作手順を示すフローチャートである。

【図 4 7】

本発明に係る標準モデル作成装置を具体的に適用したシステム例を示す図である。

【図 4 8】

クラス ID・初期標準モデル・参照モデル対応表の例を示す図である。

【図 4 9】

図 4 8 のクラス ID・初期標準モデル・参照モデル対応表における参照モデル 8 A A ~ A Z の例を示す図である。

【図 5 0】

図 4 8 のクラス ID・初期標準モデル・参照モデル対応表における参照モデル 6 4 Z A ~ Z Z の例を示す図である。

【図 5 1】

図 4 8 のクラス ID・初期標準モデル・参照モデル対応表における初期標準モデル 8 A ~ 6 4 Z の例を示す図である。

【図 5 2】

クラス ID・初期標準モデル・参照モデル対応表の作成方法を示すフローチャ

ートである。

【図 5 3】

図 5 2 におけるステップ S 1 1 0 0 の具体例を示す図である。

【図 5 4】

図 5 2 におけるステップ S 1 1 0 2 の具体例を示す図である。

【図 5 5】

図 5 2 におけるステップ S 1 1 0 3 の具体例を示す図である。

【図 5 6】

図 5 2 におけるステップ S 1 1 0 4 の具体例を示す図である。

【図 5 7】

端末がサーバと通信することによってクラス ID・初期標準モデル・参照モデル対応表を完成させる手順を示す図である。

【図 5 8】

クラス ID・初期標準モデル・参照モデル対応表を用いた初期標準モデルの決定方法を示すフローチャートである。

【図 5 9】

図 5 8 におけるステップ S 1 1 0 5 の具体例を示す図である。

【図 6 0】

第 3 近似部を用いて作成した標準モデルを用いた認識実験の結果を示すグラフである。

【符号の説明】

1 0 1、4 0 1、5 0 1、6 0 1、7 0 1、8 0 1 サーバ
1 0 2、2 0 2、3 0 2、4 0 2、5 0 2、6 0 2、7 0 2、9 0 2 参
照モデル準備部
1 0 3、2 0 3、3 0 3、4 0 3、5 0 3、6 0 3、7 0 3、9 0 3、1 0
0 3 参照モデル記憶部
1 0 4、2 0 6、3 0 6、4 0 6、5 0 6、6 0 6、7 0 6、9 0 6、1 0
0 6 標準モデル作成部
1 0 4 a、2 0 6 a、3 0 6 a、9 0 6 a、1 0 0 6 a 標準モデル構造決

定部

104b、206b、306b、906b、1006b 初期標準モデル作

成部

104c、206c、306c、906c 統計量記憶部

104d、206d、306d、906d 統計量推定部

104e 第1近似部

111、311、511、611、711 読み込み部

112、413、513 書き込み部

201 STB

204、304、904 利用情報作成部

205、305、405、505、605、705、905 参照モデル

選択部

206e 一般近似部

211、312、912、1012 マイク

212、512 音声データ蓄積部

213、913、1013 音声認識部

301、1001 PDA

306e 第2近似部

307、607、907 仕様情報作成部

313 雑音識別部

404、504、604、704 利用情報受信部

411 カメラ

412 画像データ蓄積部

414 電話

507、707 仕様情報受信部

512、614、712 端末装置

708 標準モデル記憶部

709 標準モデル送信部

712a、901 携帯電話機

712b カーナビゲーション装置

810、909 参照モデル受信部

906e 第3近似部

908 類似度情報作成部

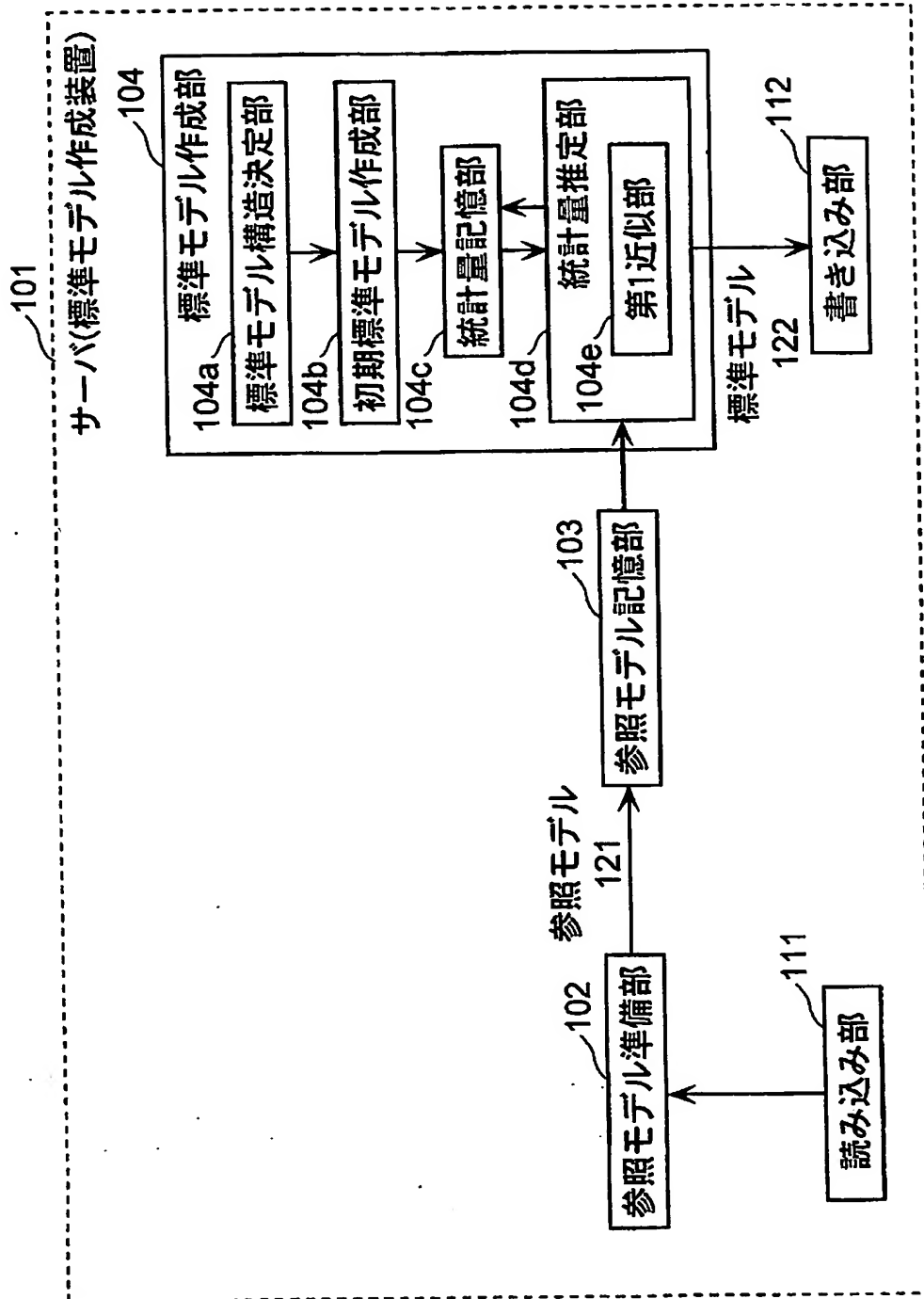
910 サーバ装置

1014 アプリ・仕様情報対応データベース

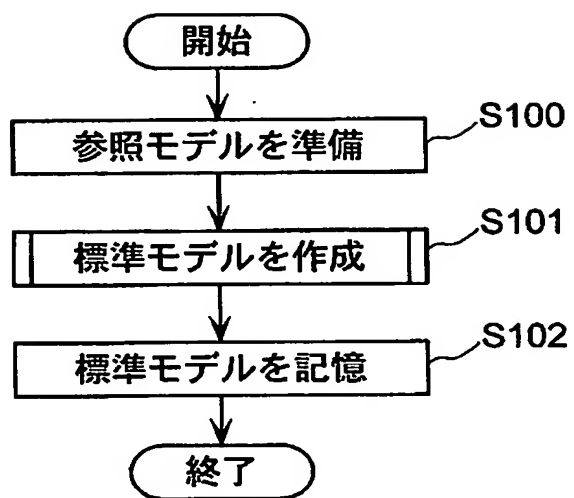
【書類名】

図面

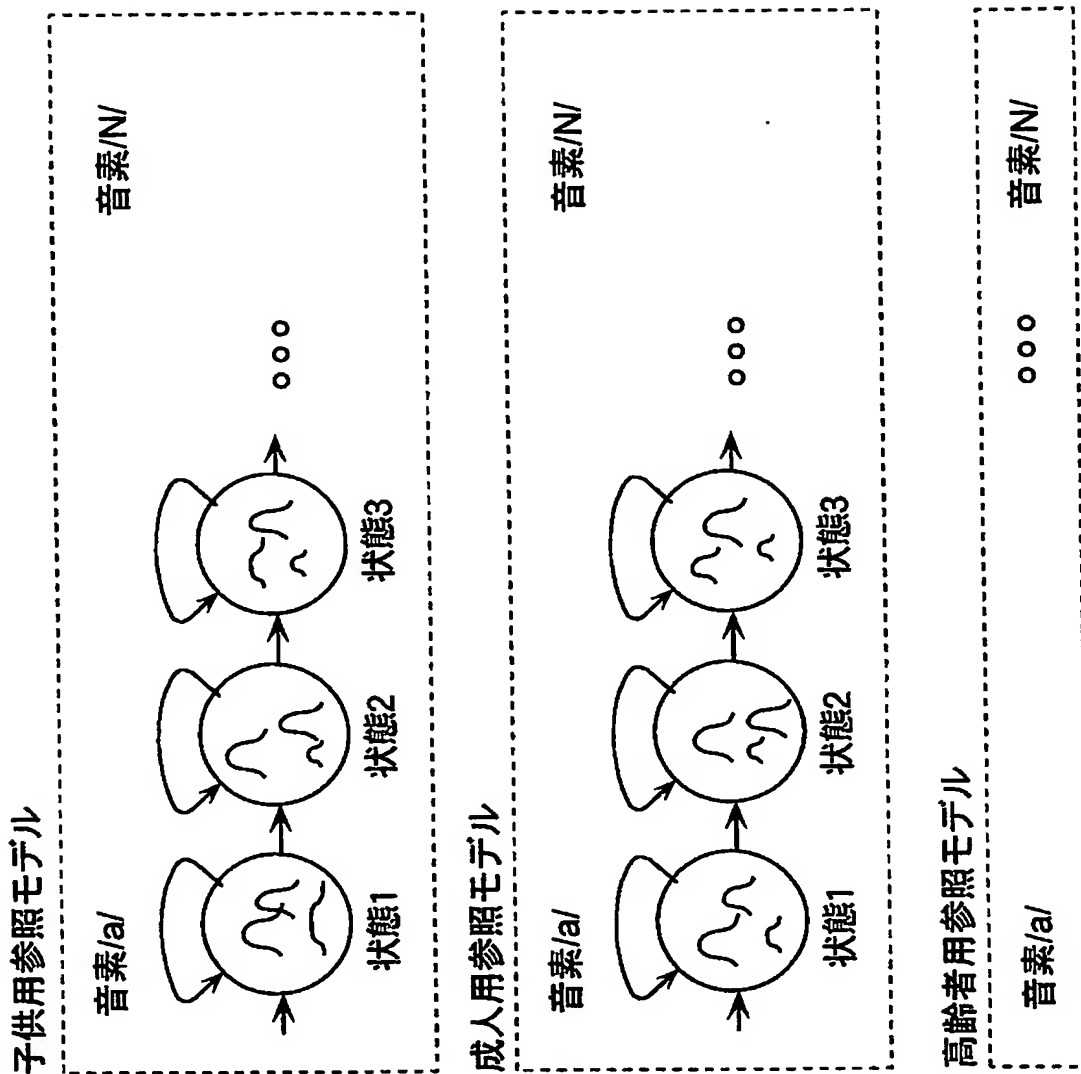
【図 1】



【図2】

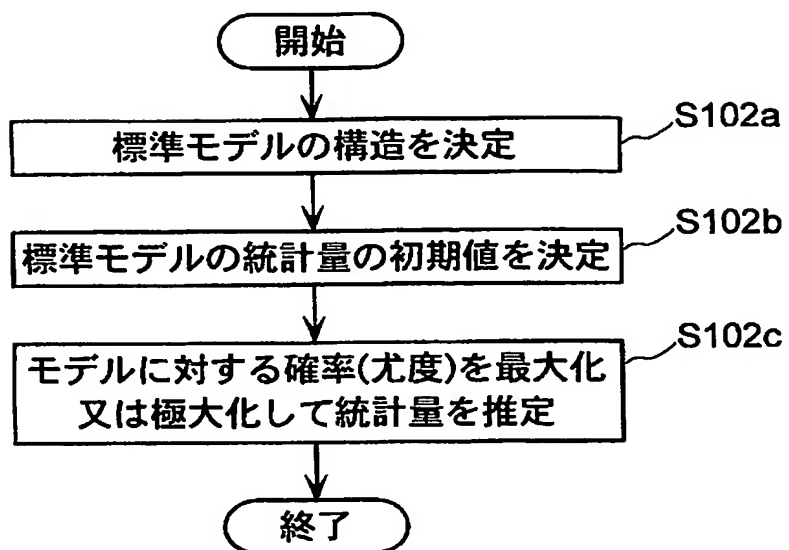


【図 3】

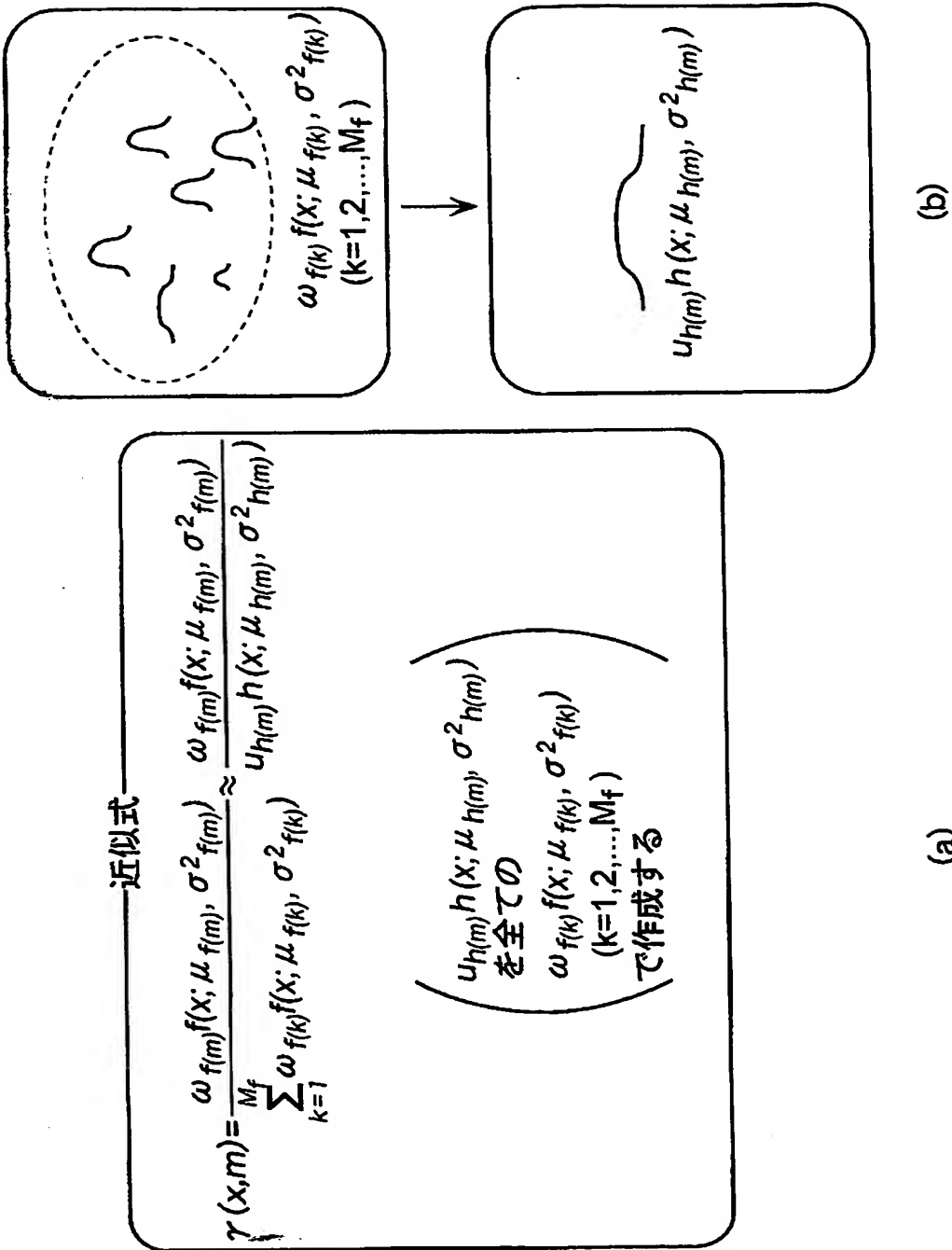


参照モデル
121

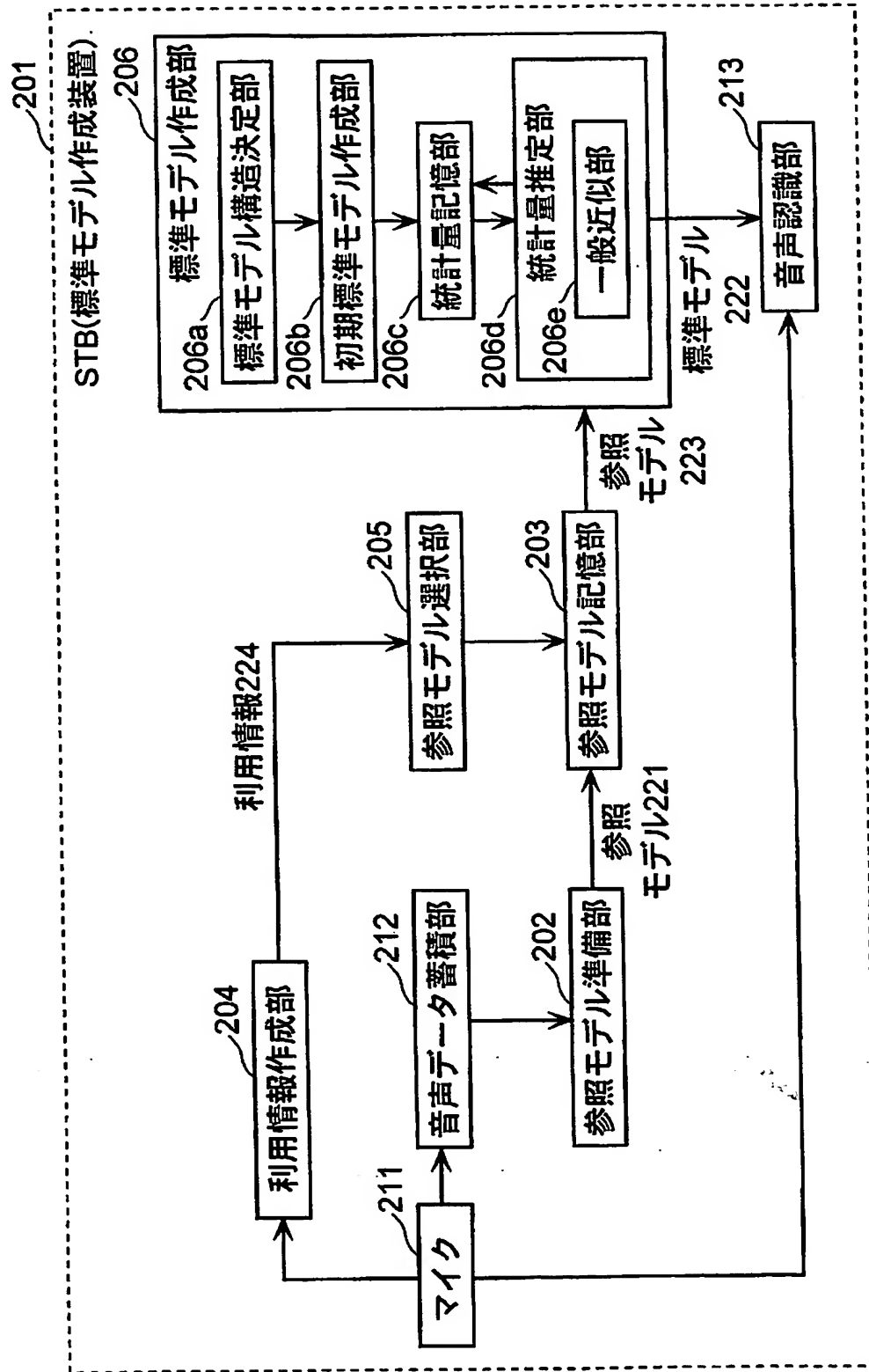
【図 4】



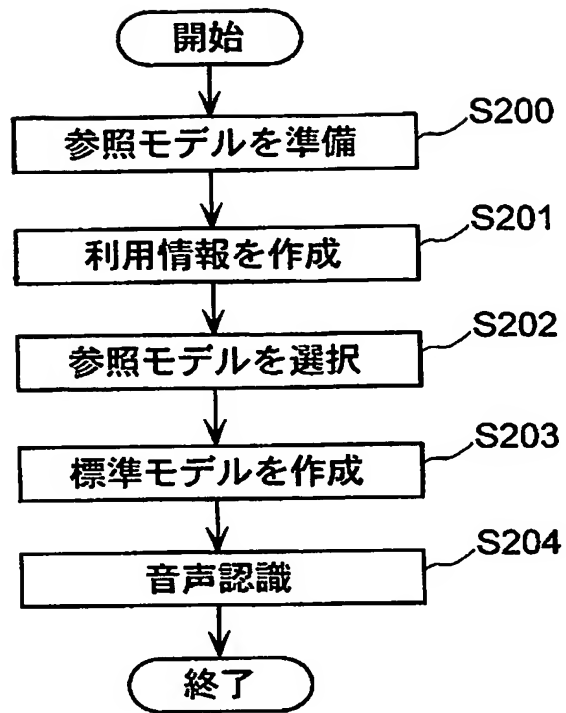
【図 5】



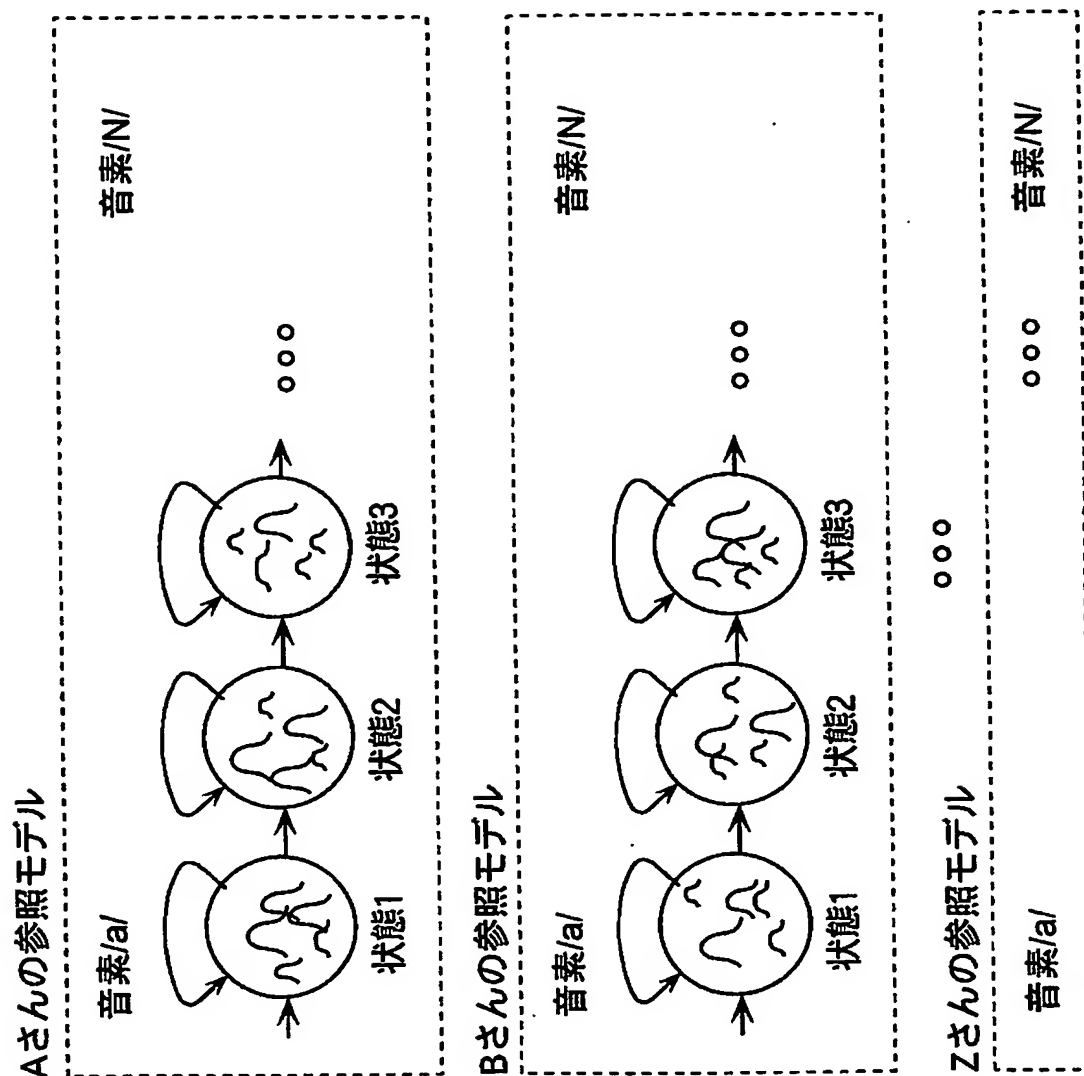
【図 6】



【図 7】

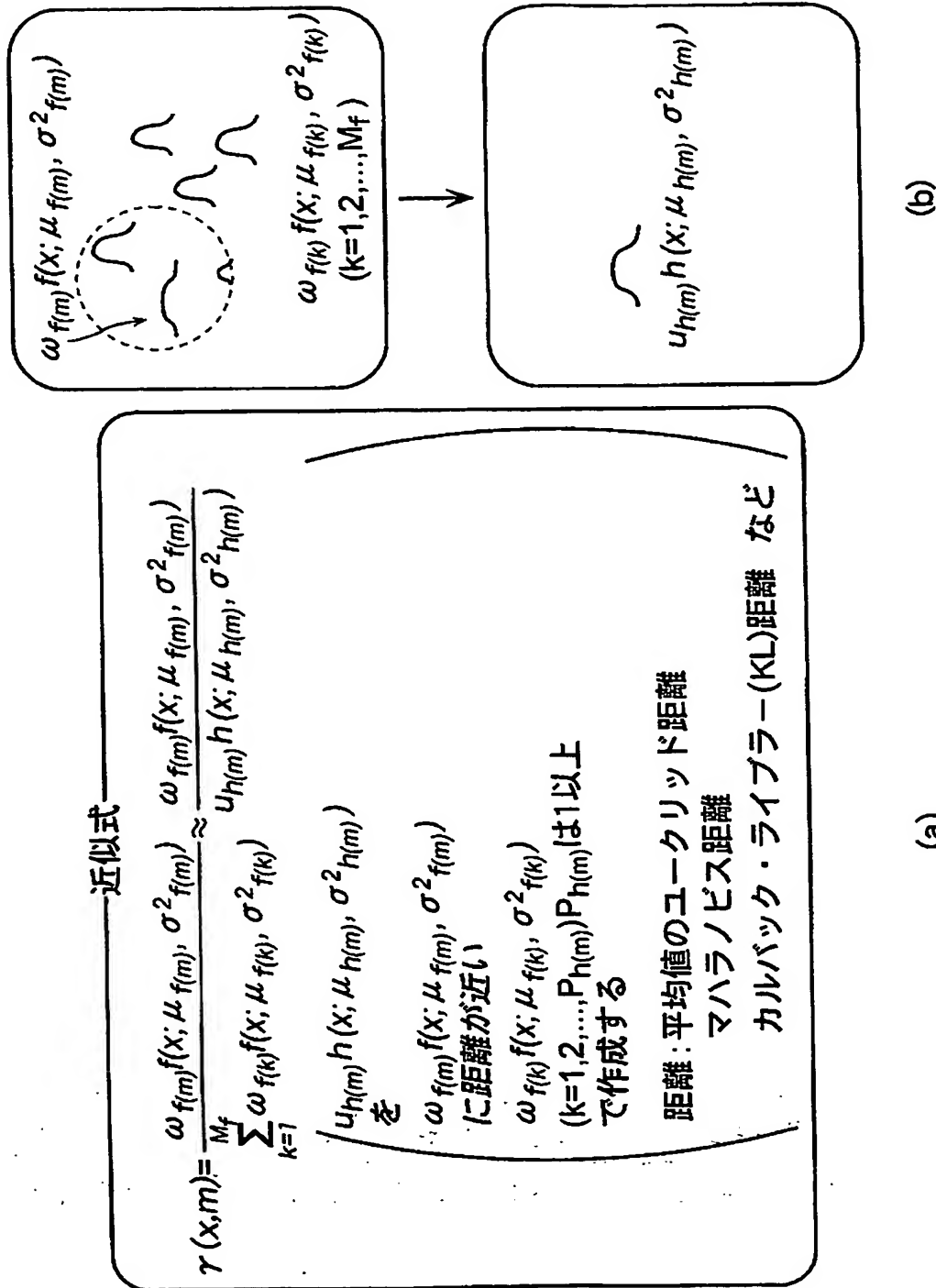


【図 8】

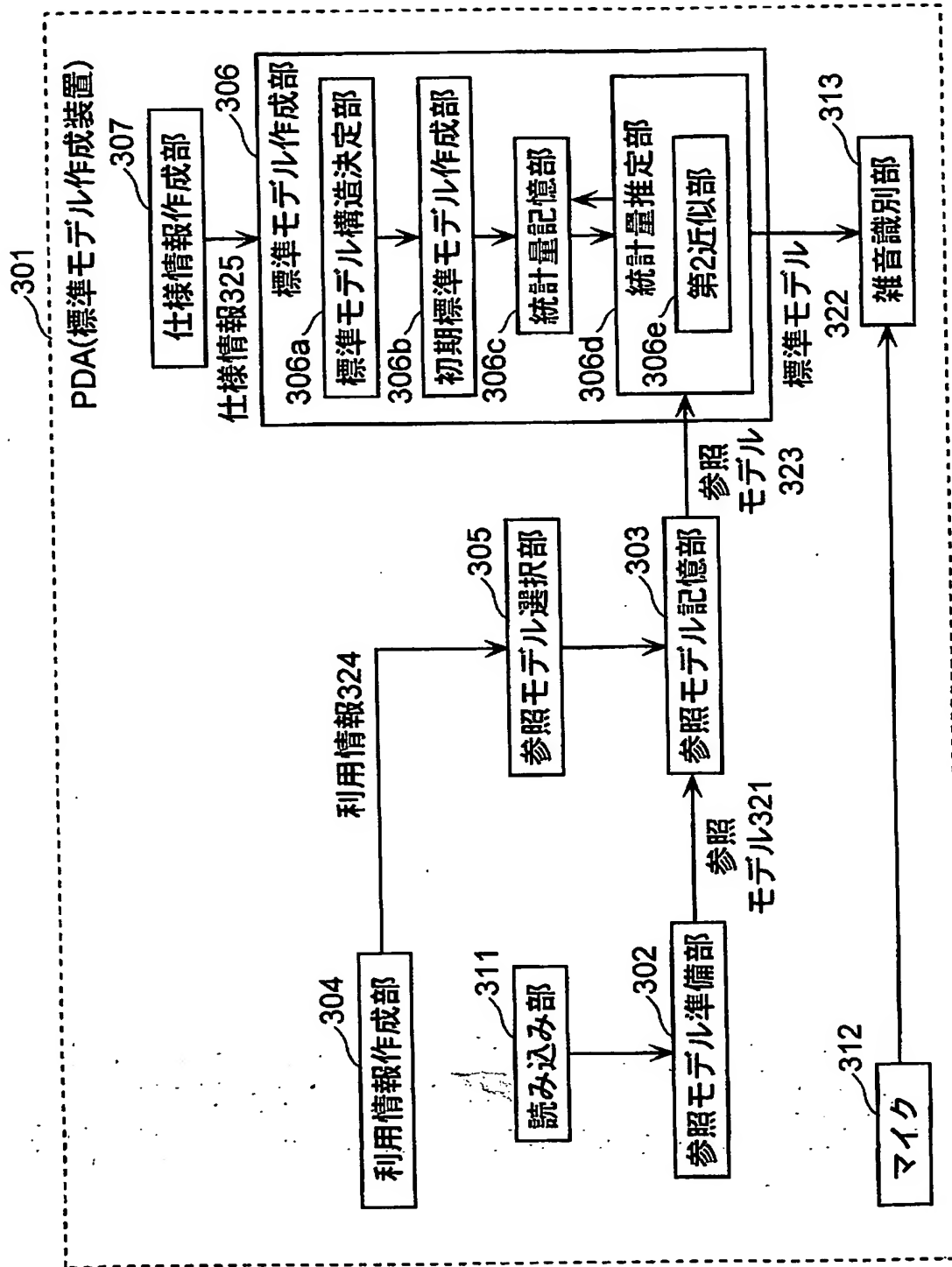


参照モデル
221

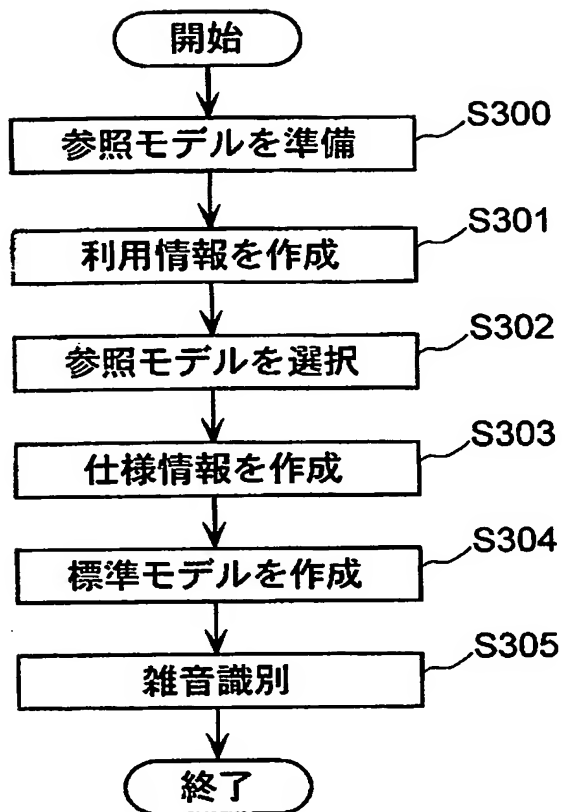
【図 9】



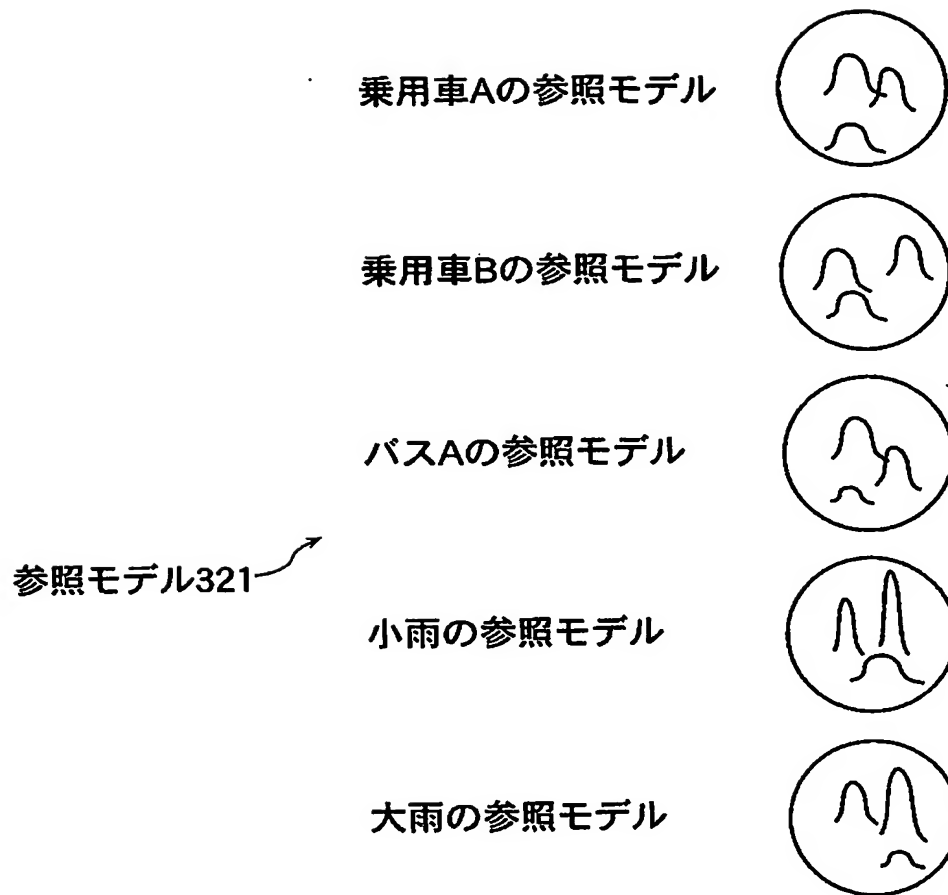
【図10】



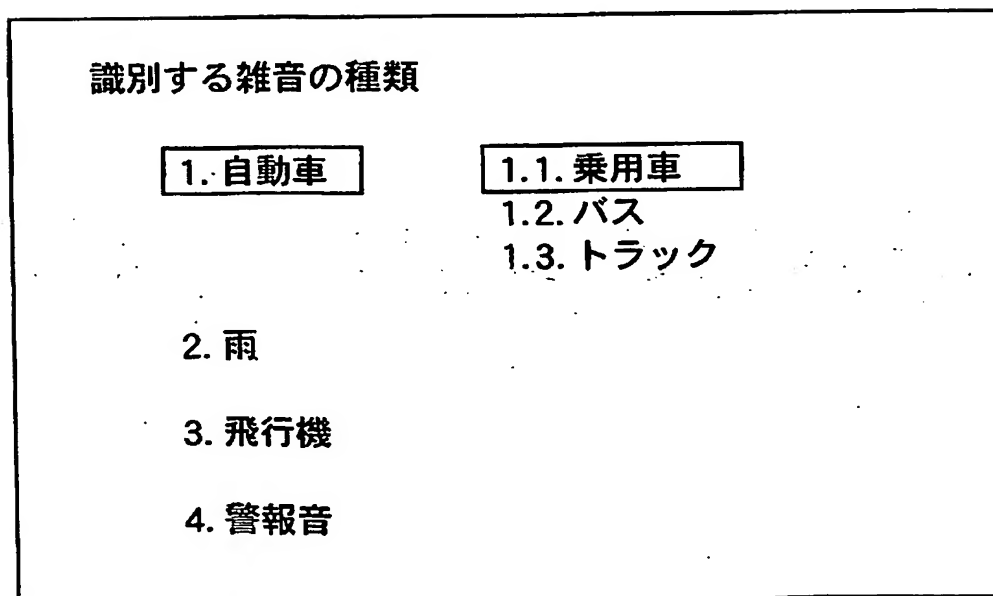
【図 11】



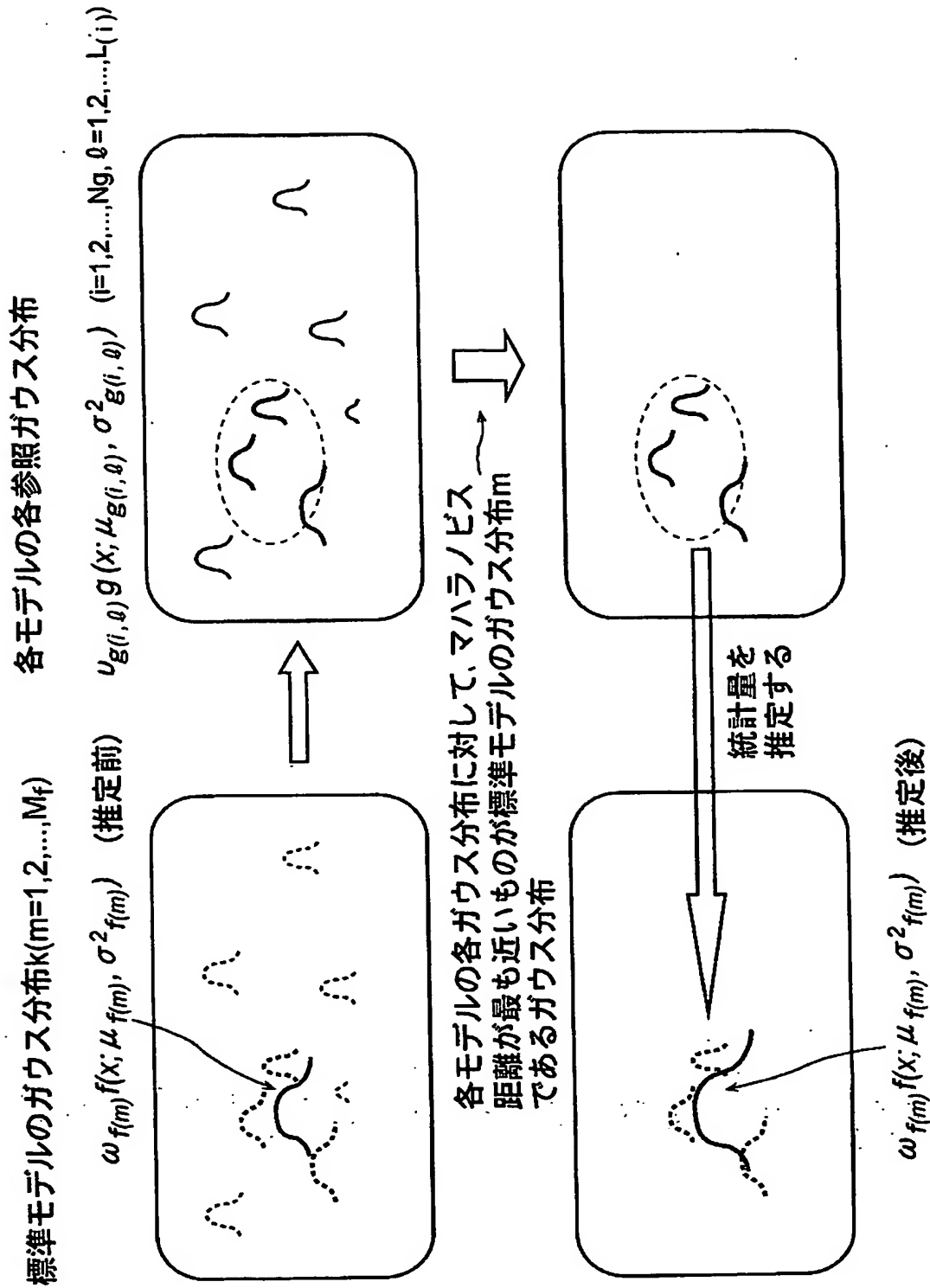
【図 12】



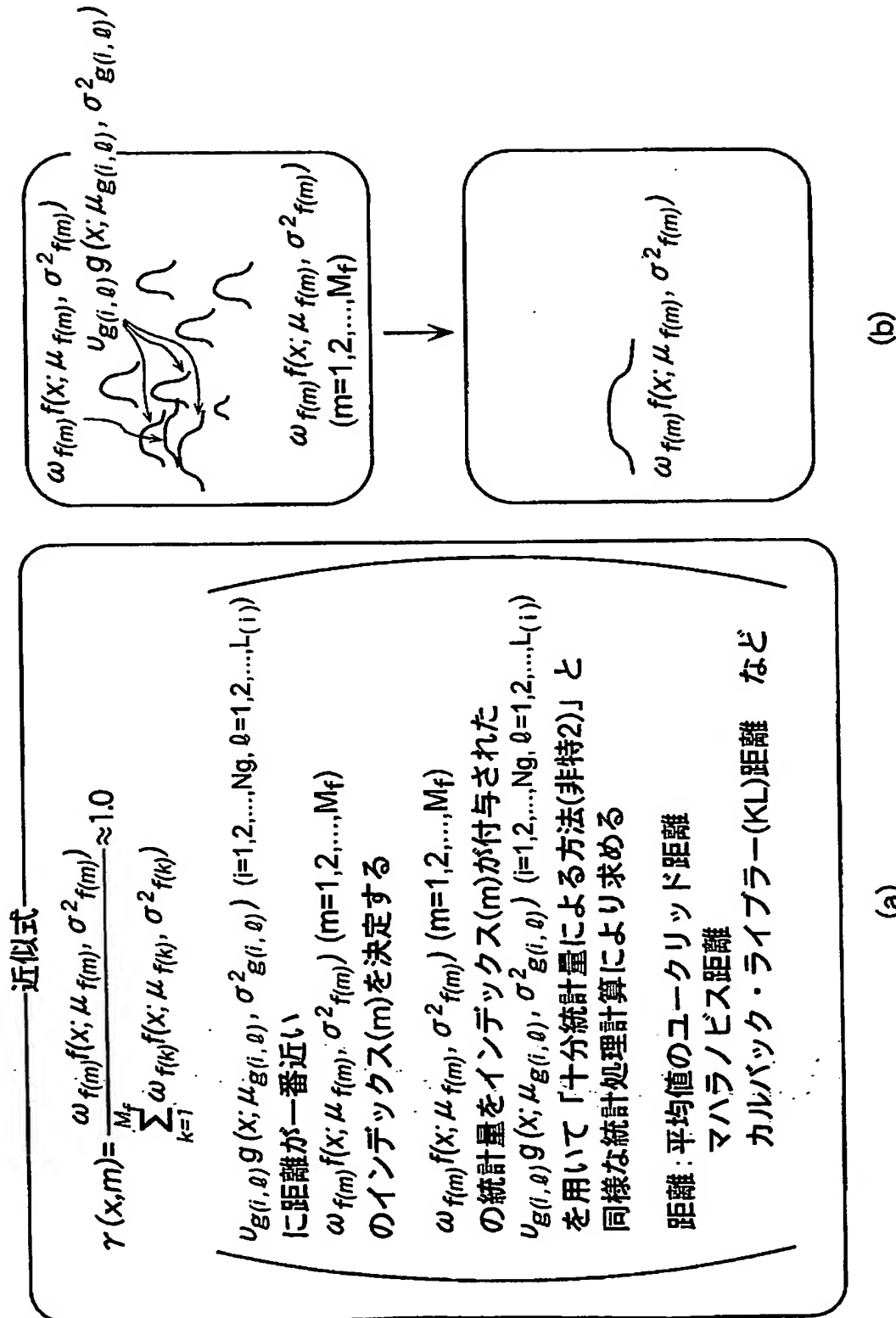
【図 13】



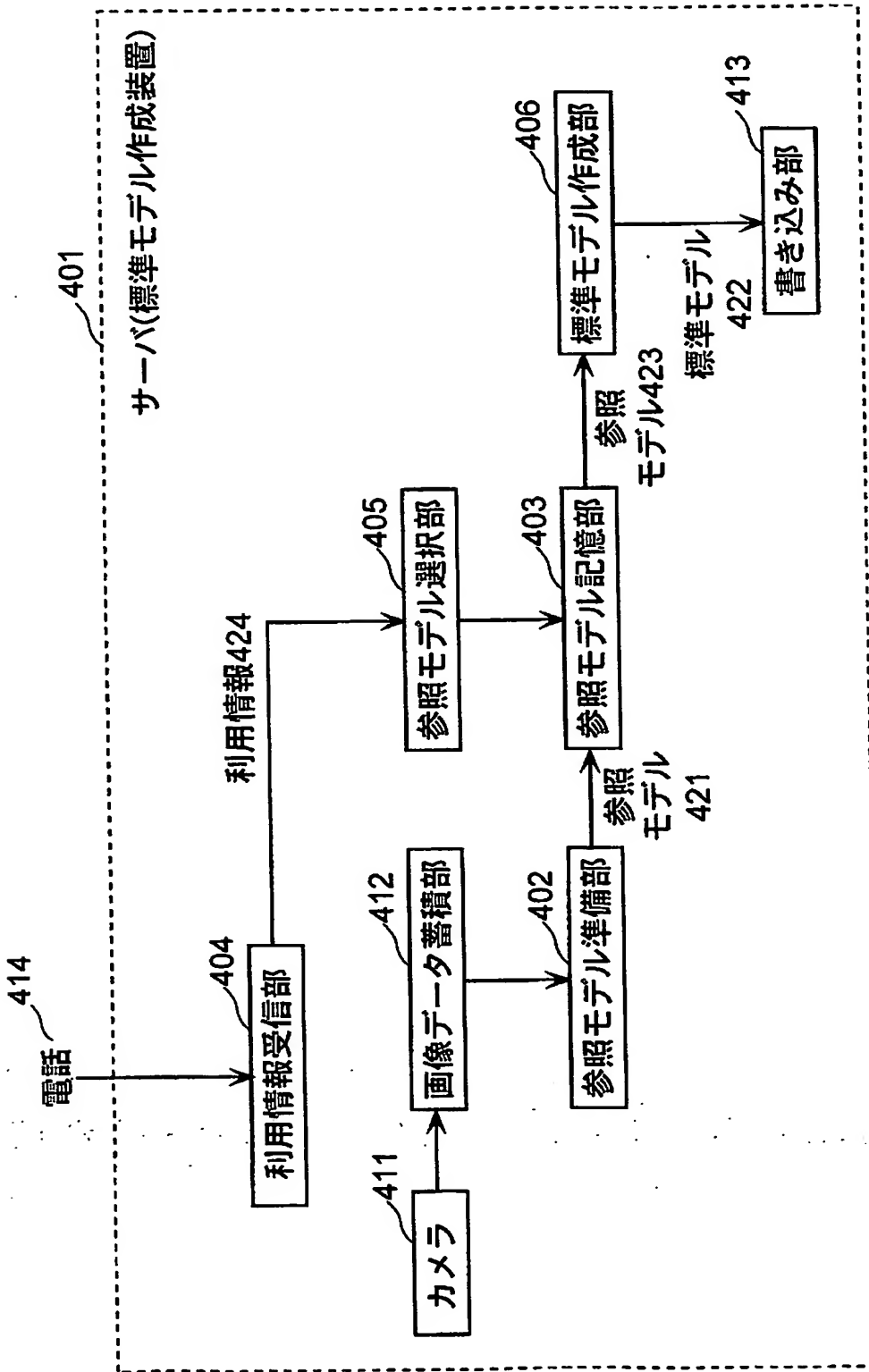
【図 14】



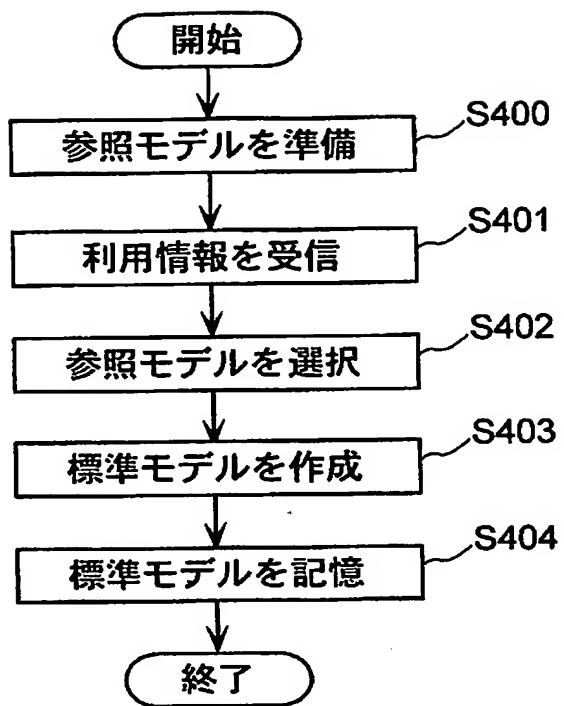
【図 15】



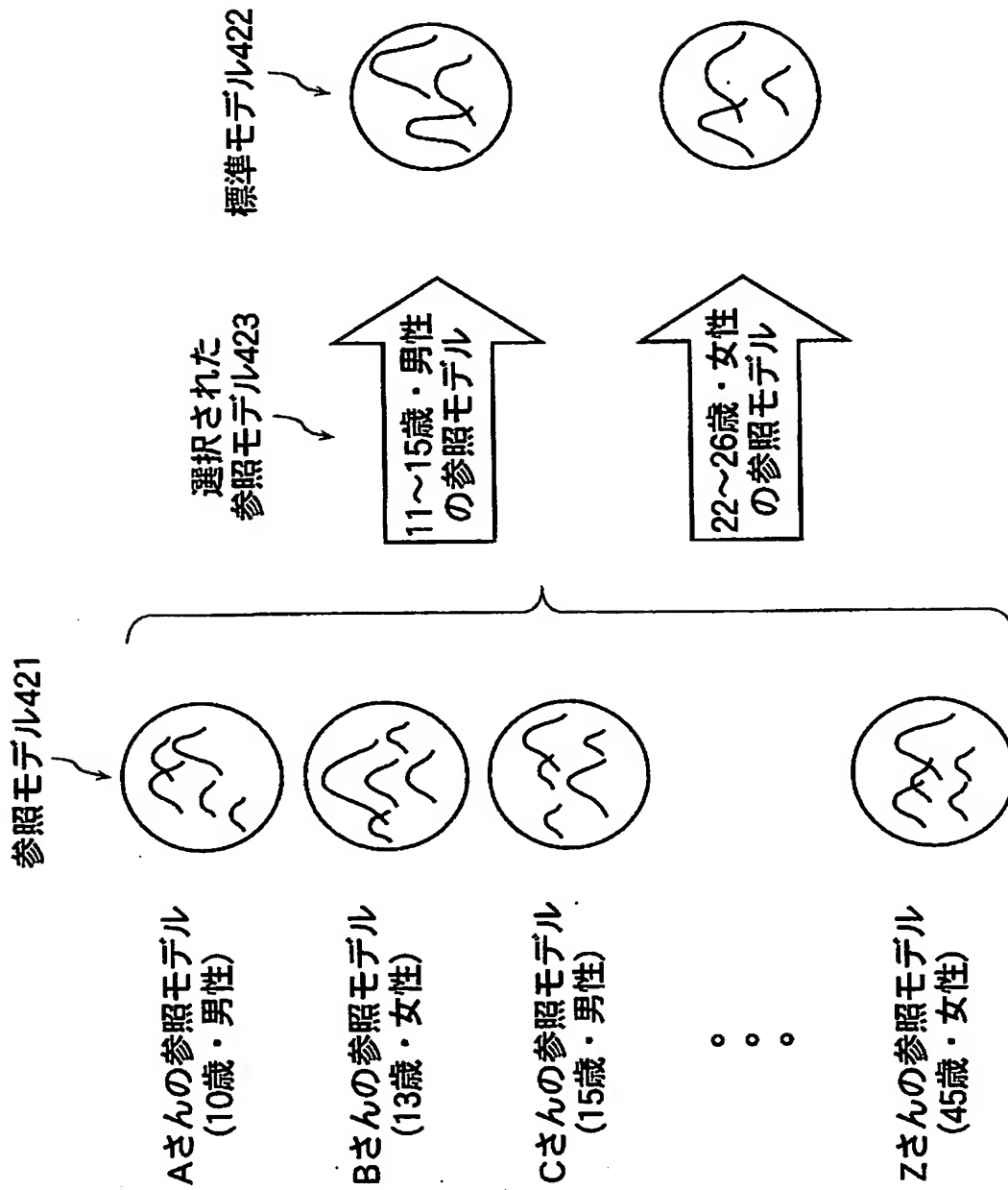
【図16】



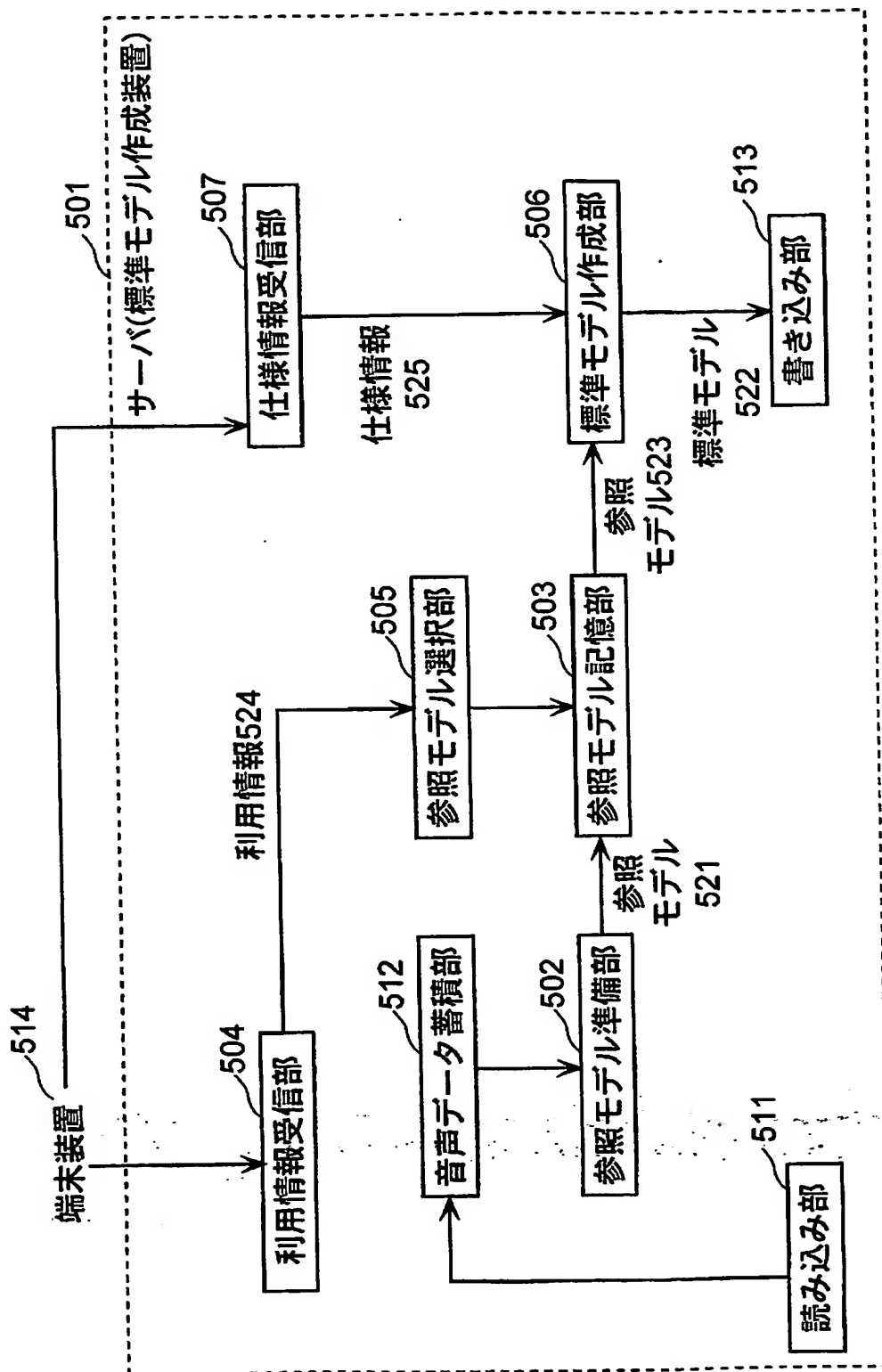
【図 17】



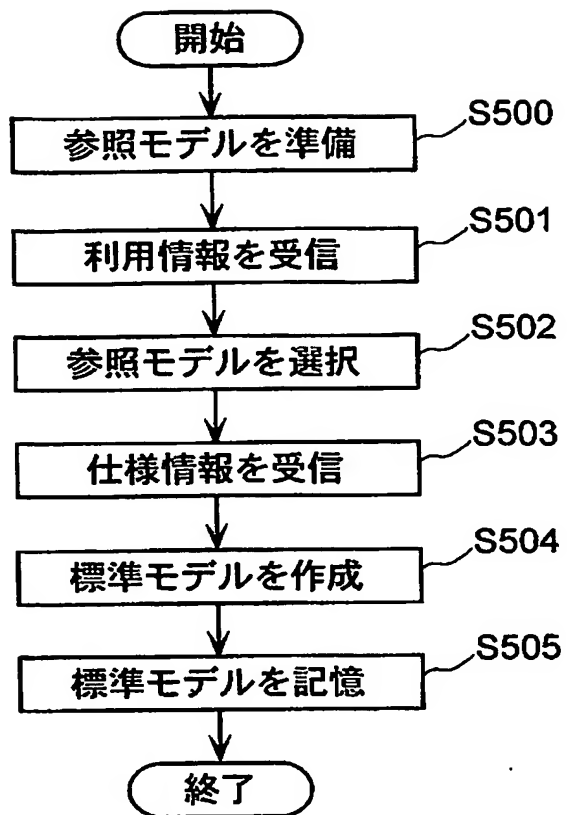
【図18】



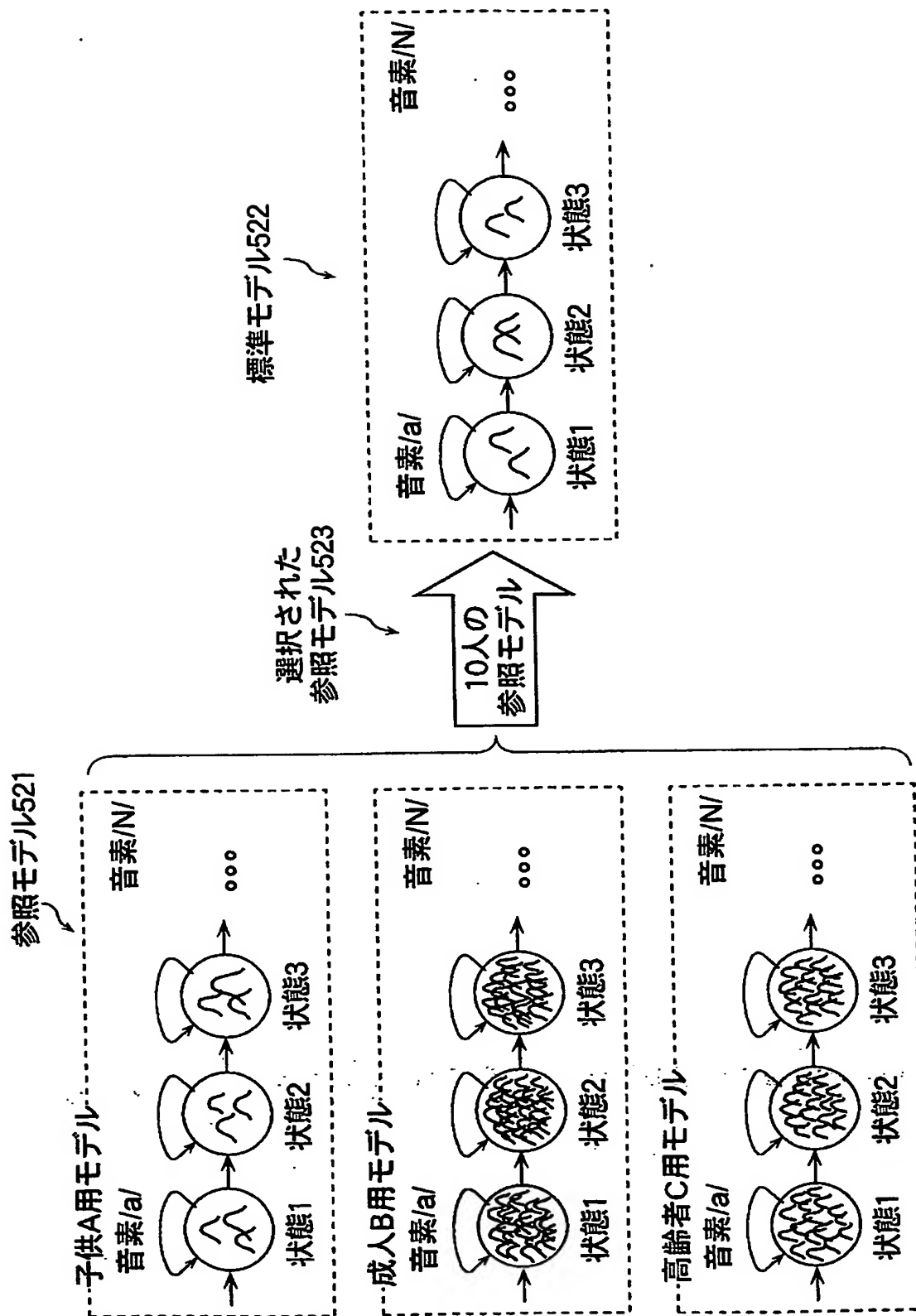
【図 19】



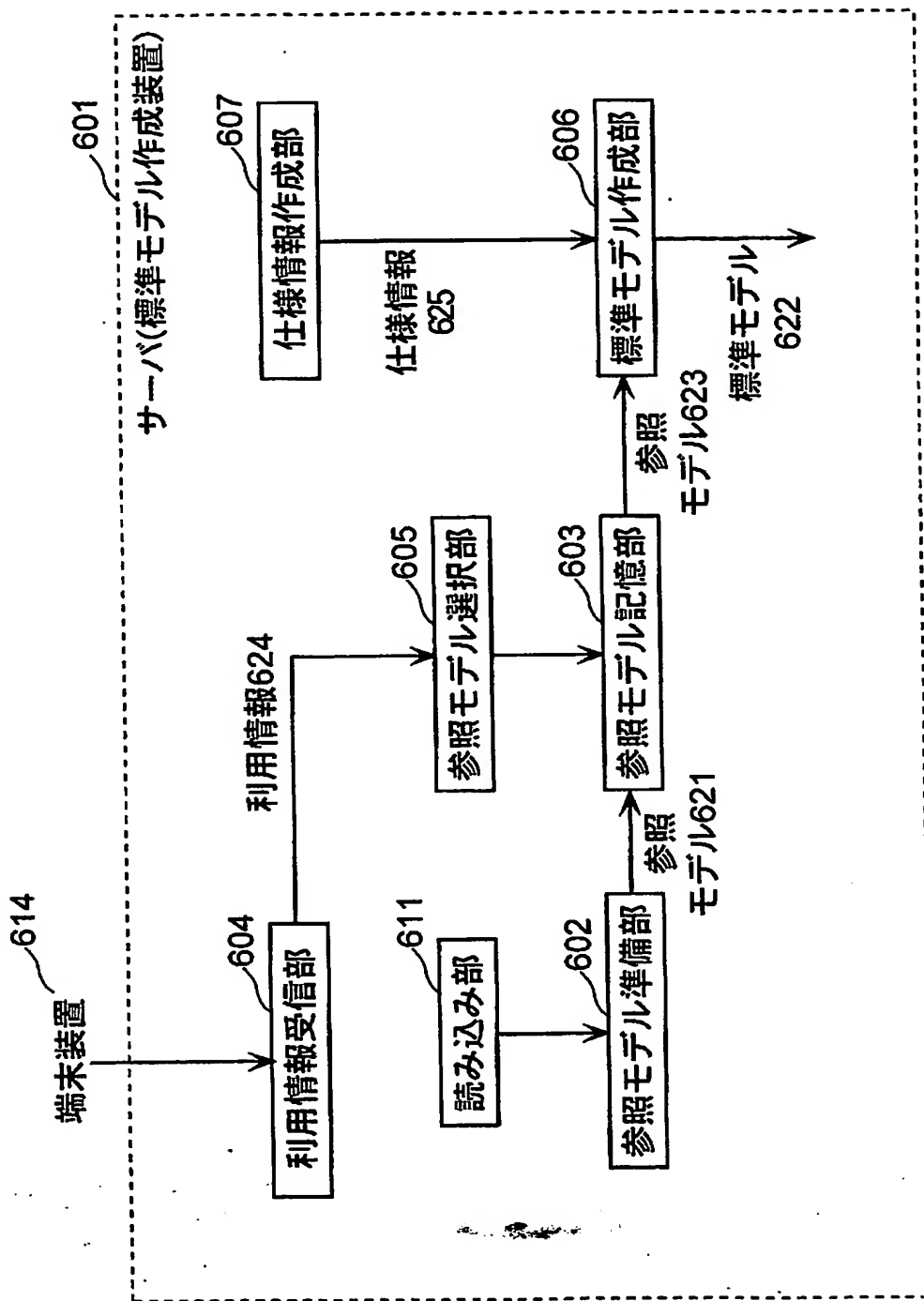
【図 20】



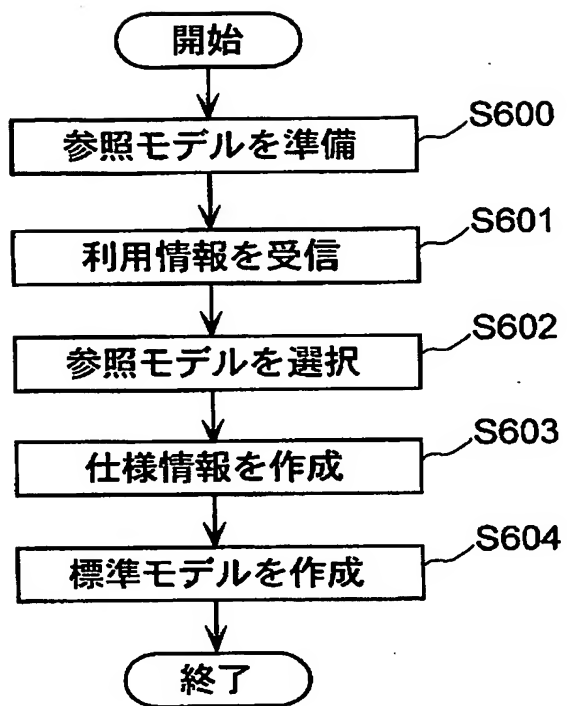
【図 21】



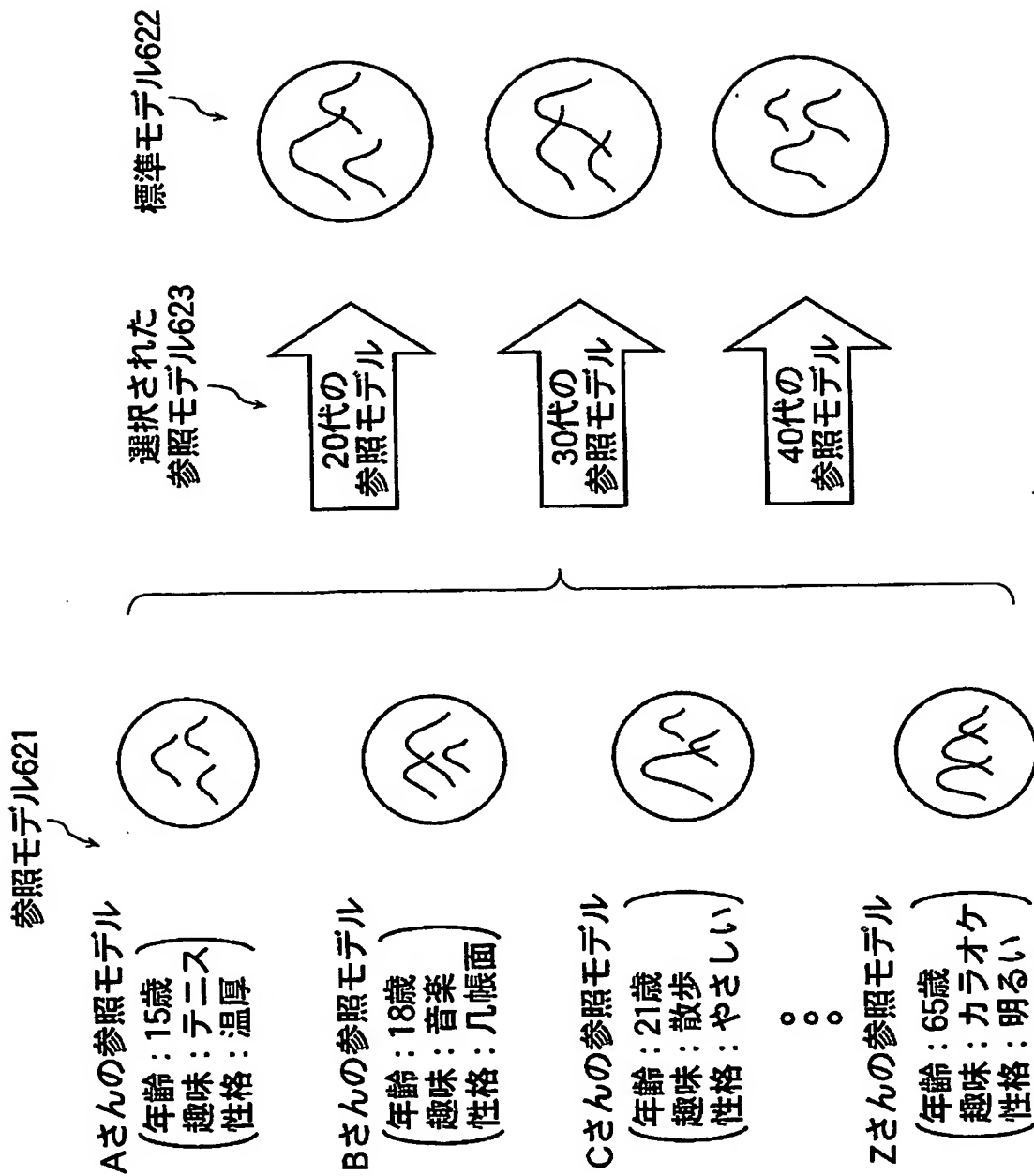
【図 2 2】



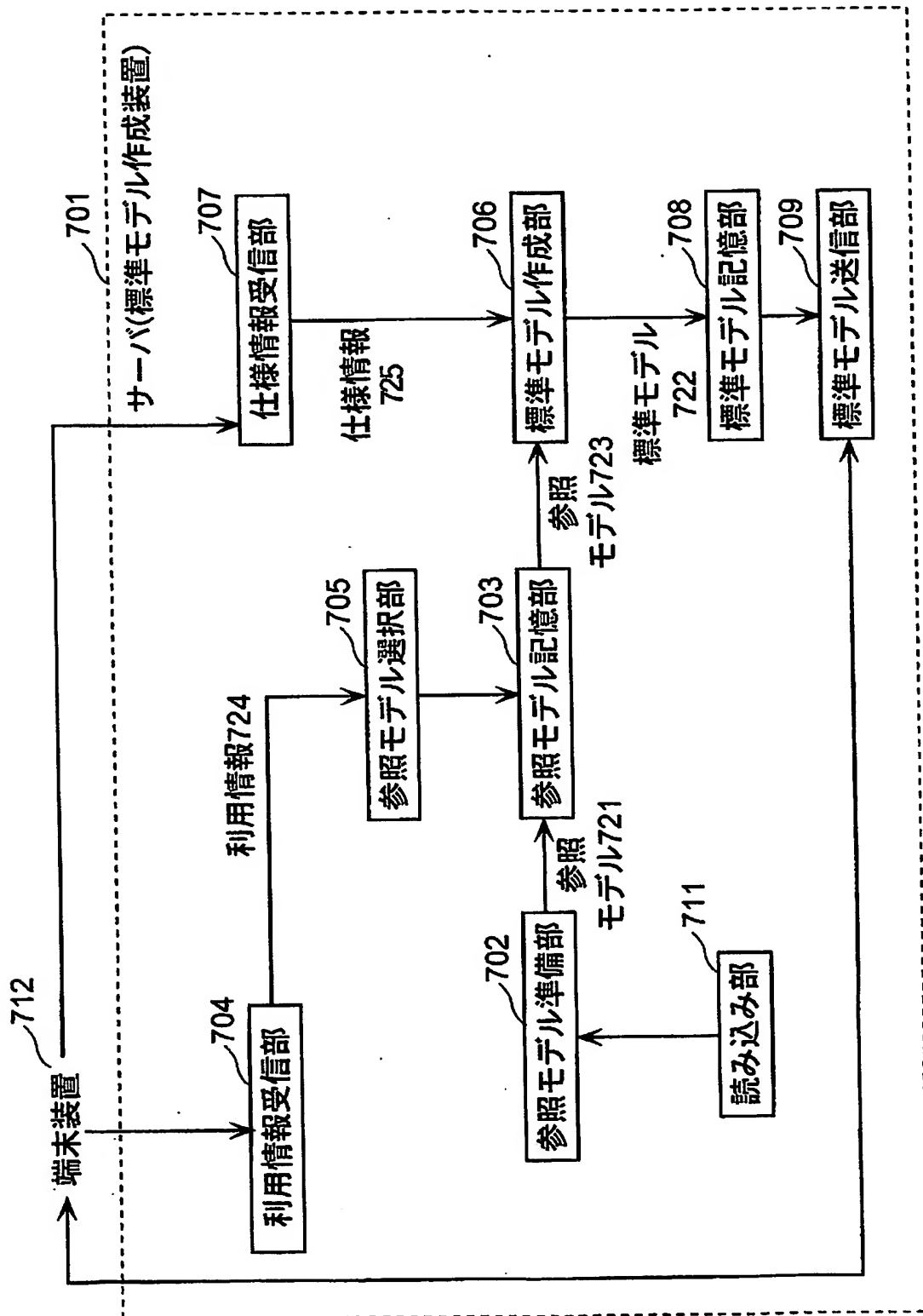
【図 23】



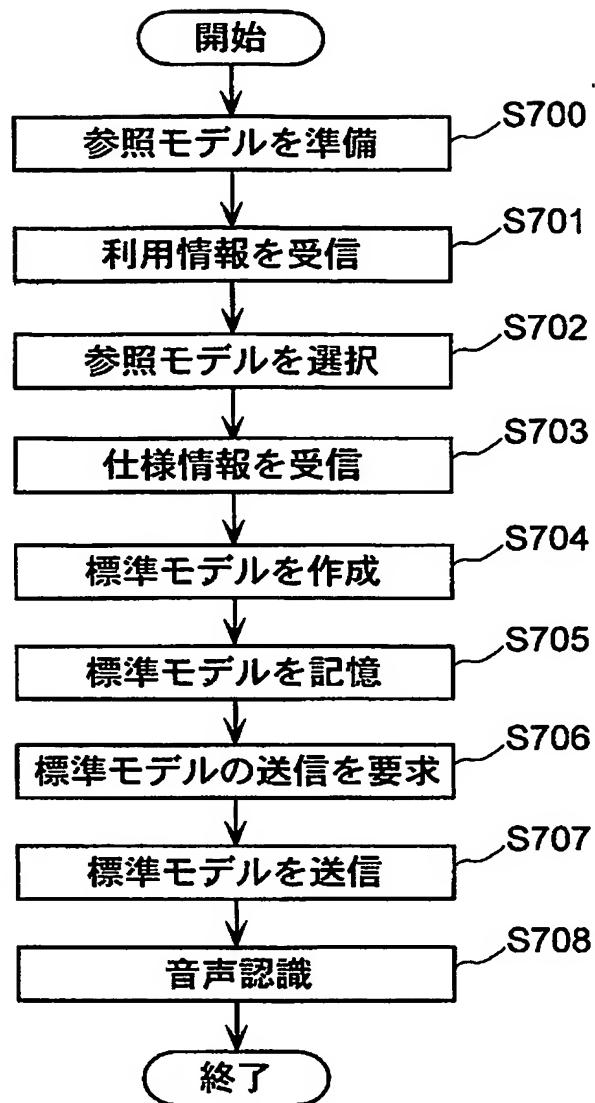
【図 24】



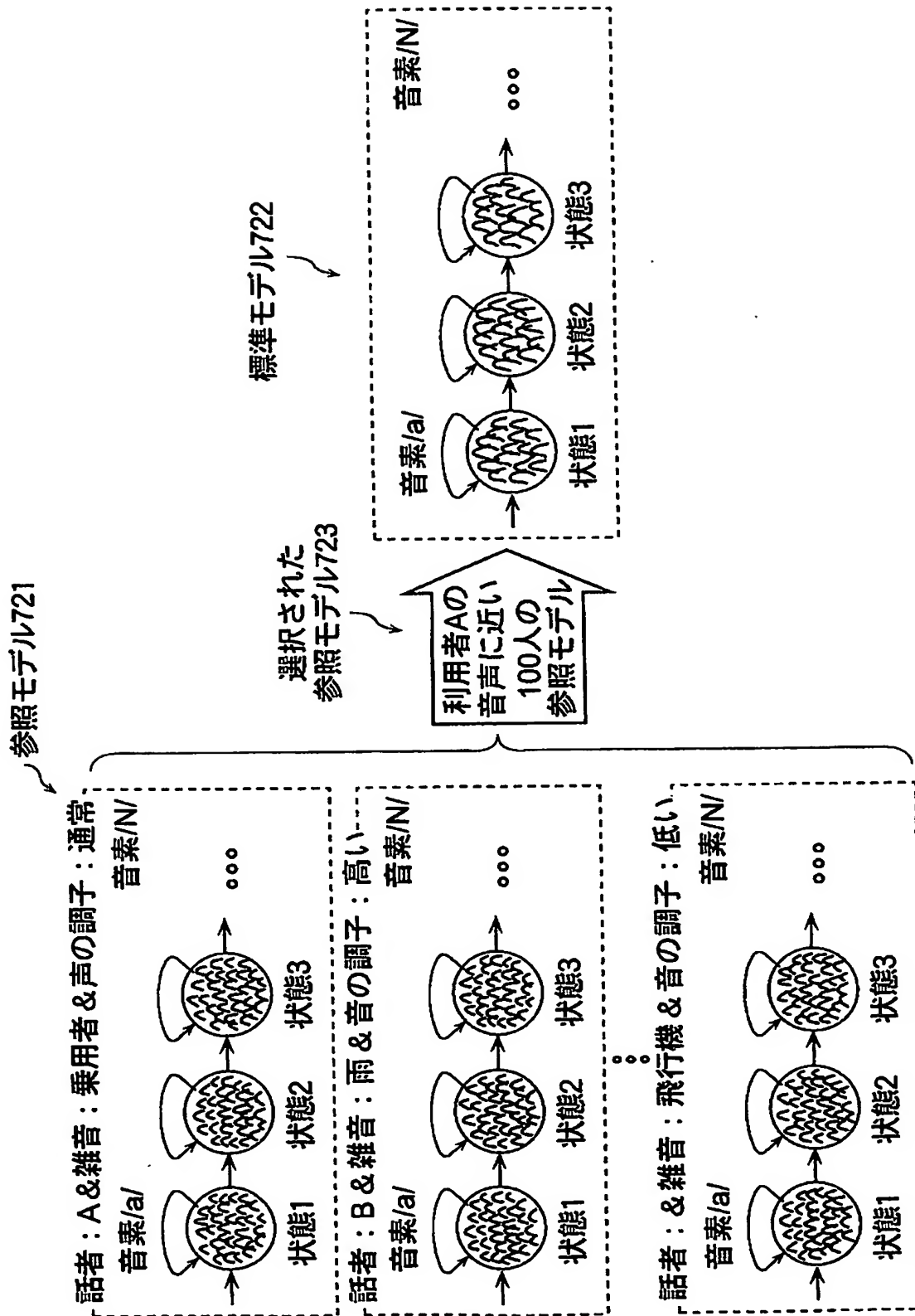
【図 25】



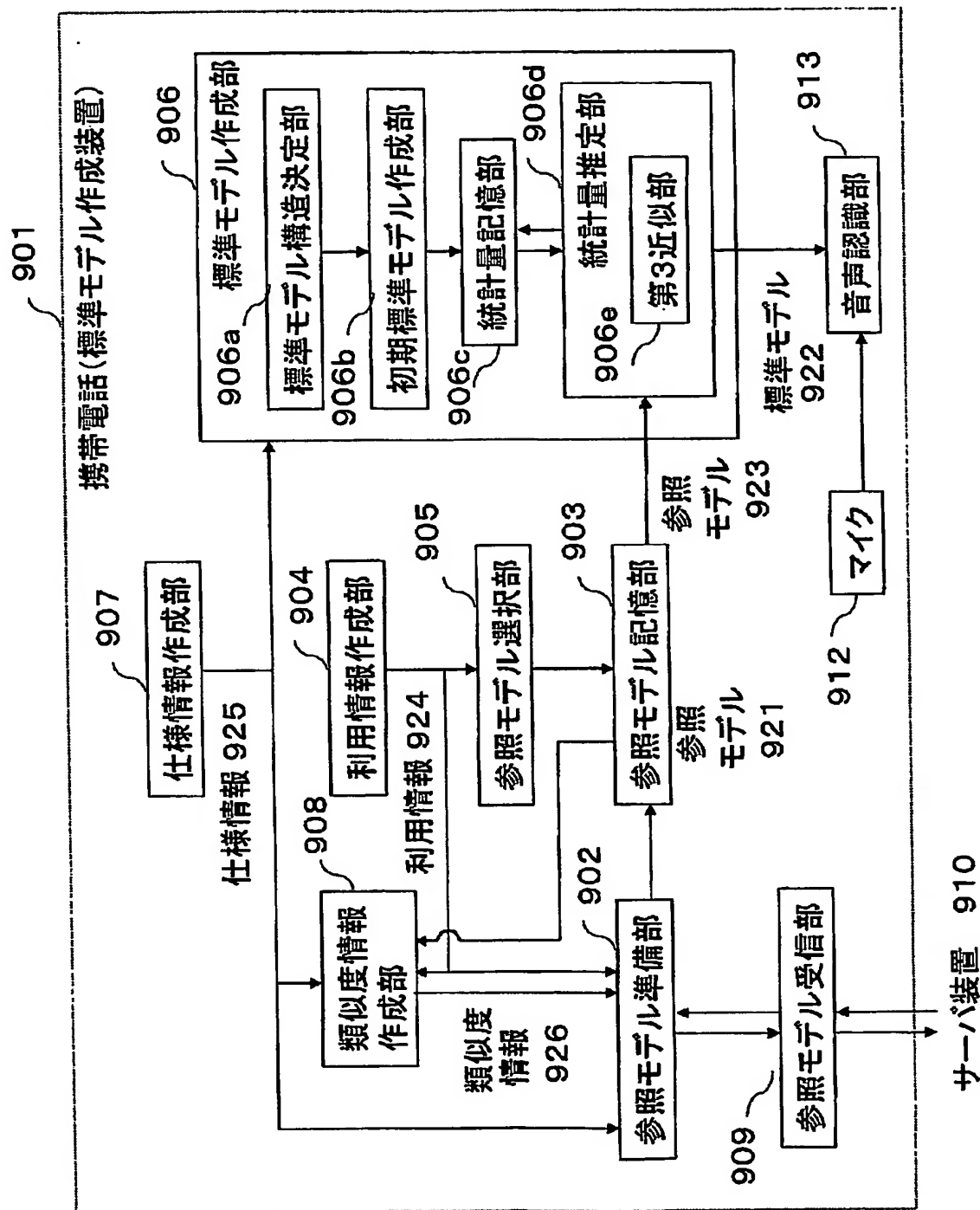
【図 26】



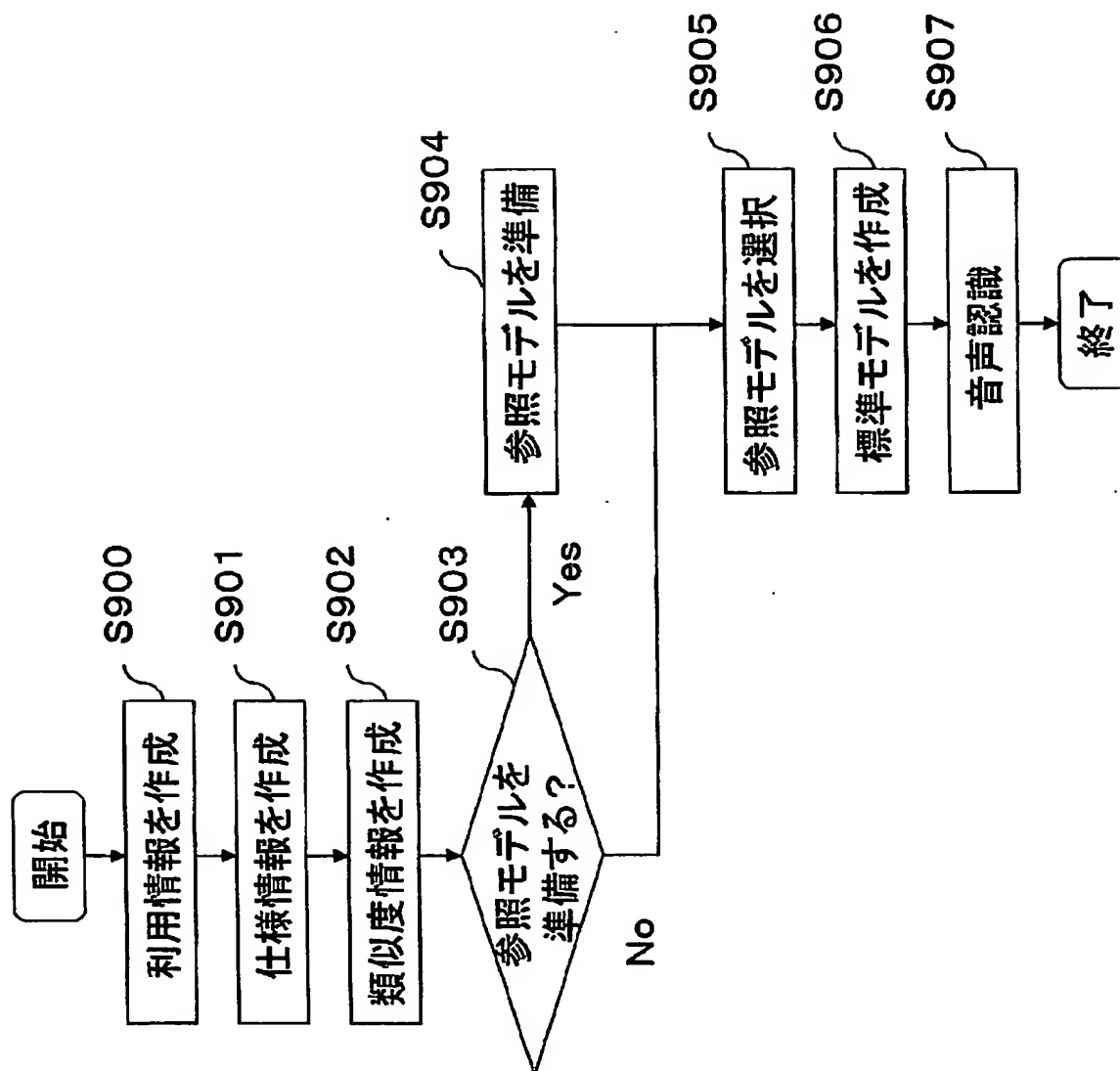
【図 27】



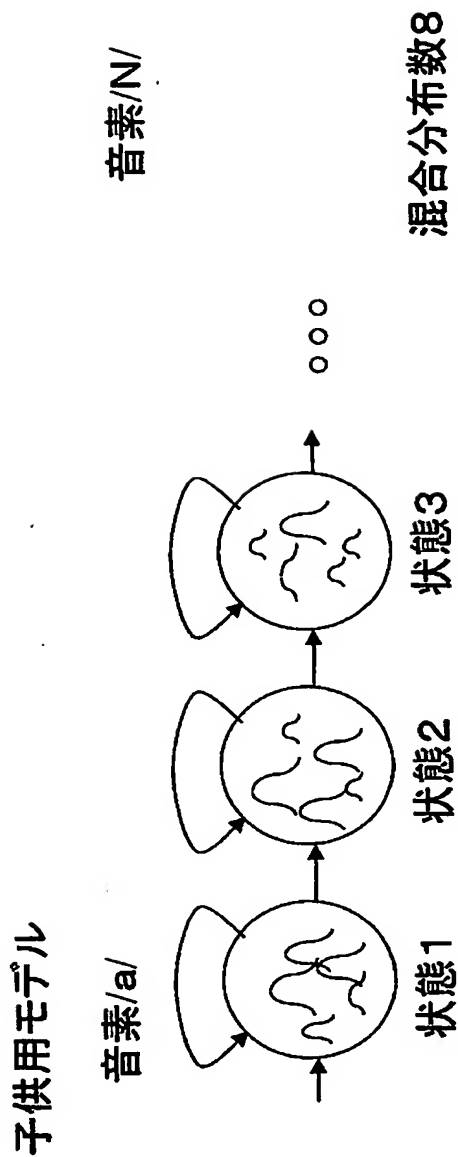
【図 28】



【図 29】



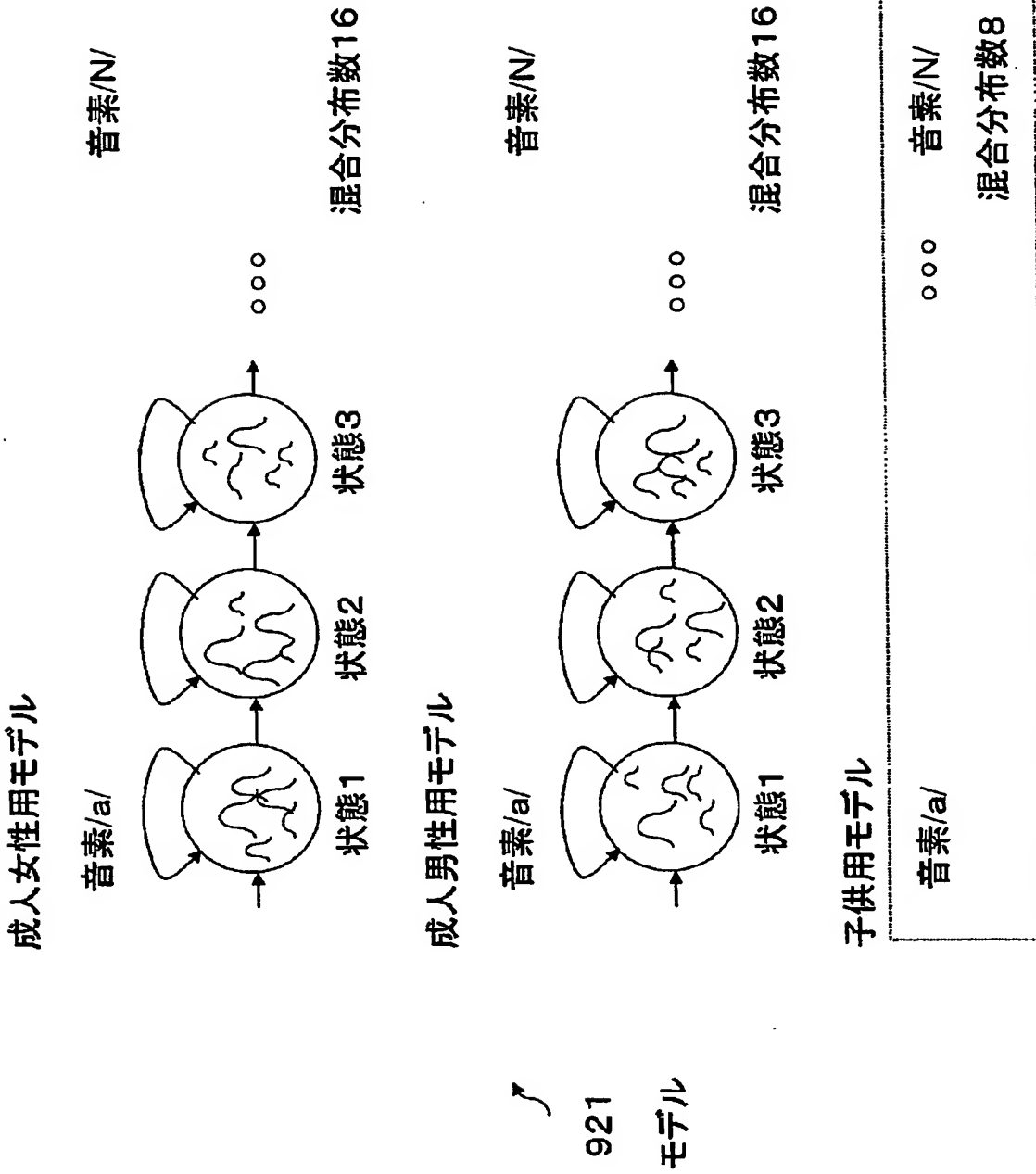
【図 30】



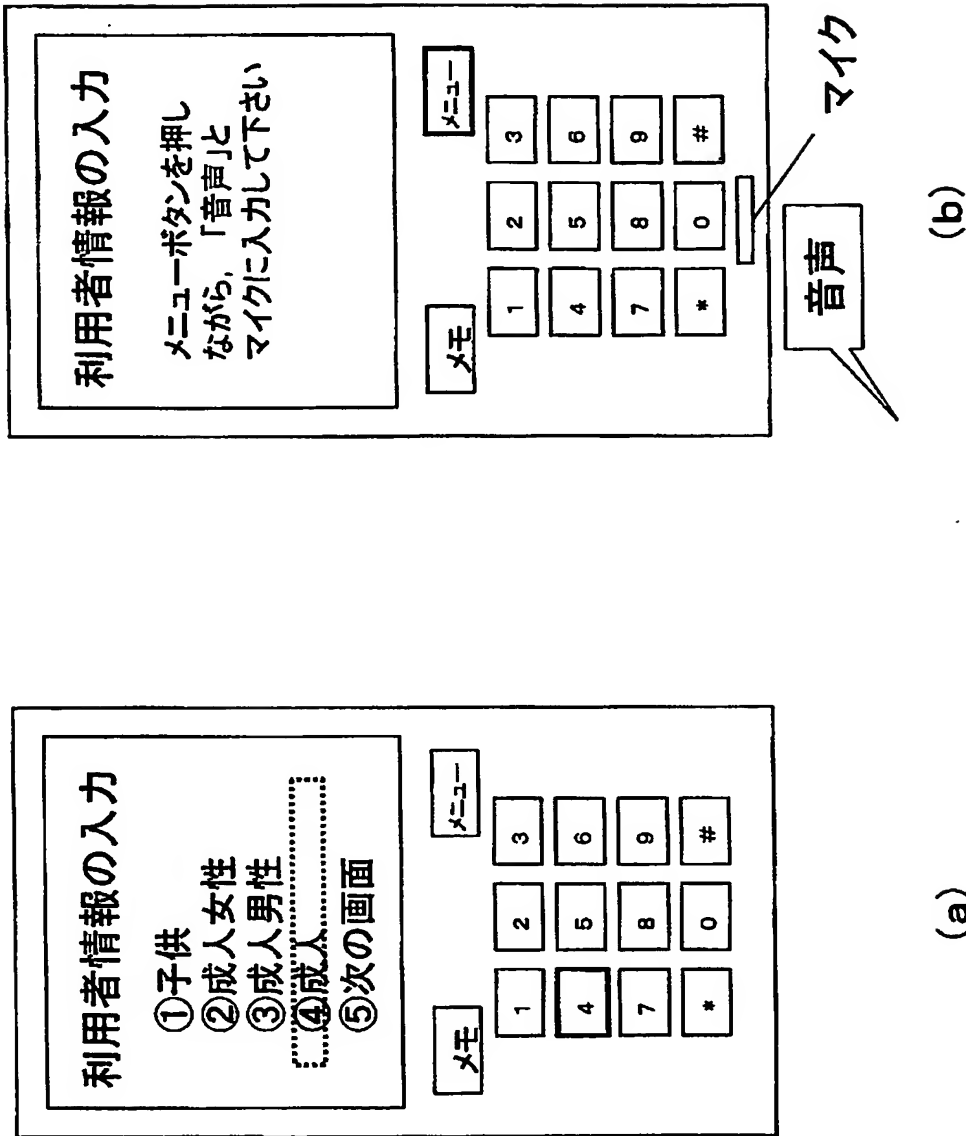
921

モデル

【図 31】



【図 3 2】



【図 33】

参照モデルをダウン
ロードします

...

完了

メモ

メニュー

1	2	3	4	5	6	7	8	9	*	0	#
---	---	---	---	---	---	---	---	---	---	---	---

(b)

参照モデルをダウン
ロードしますか？

Yes
⇒メモ

No
⇒メニュー

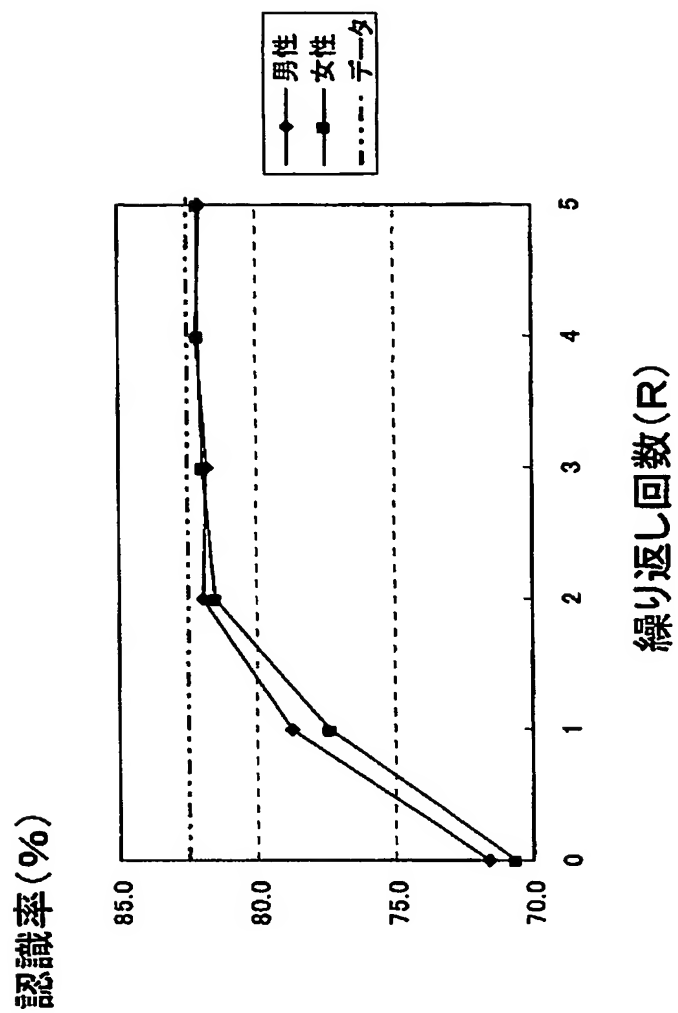
メモ

メニュー

1	2	3	4	5	6	7	8	9	*	0	#
---	---	---	---	---	---	---	---	---	---	---	---

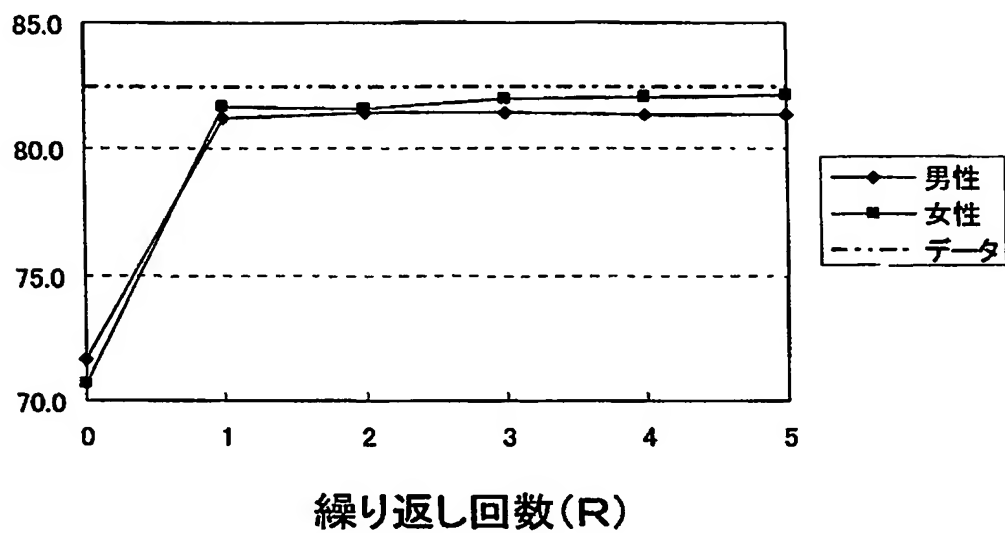
(a)

【図 34】

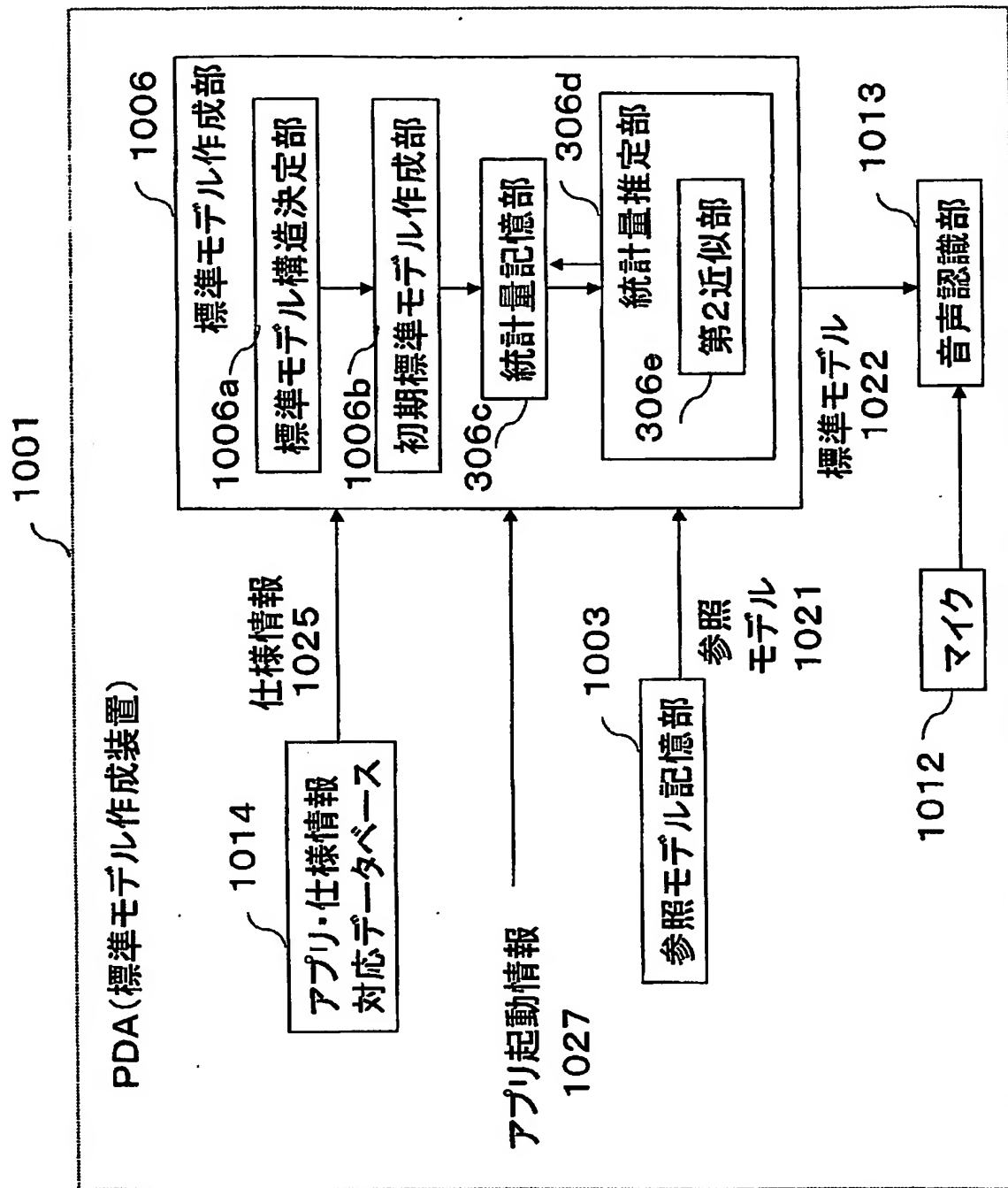


【図 35】

認識率 (%)



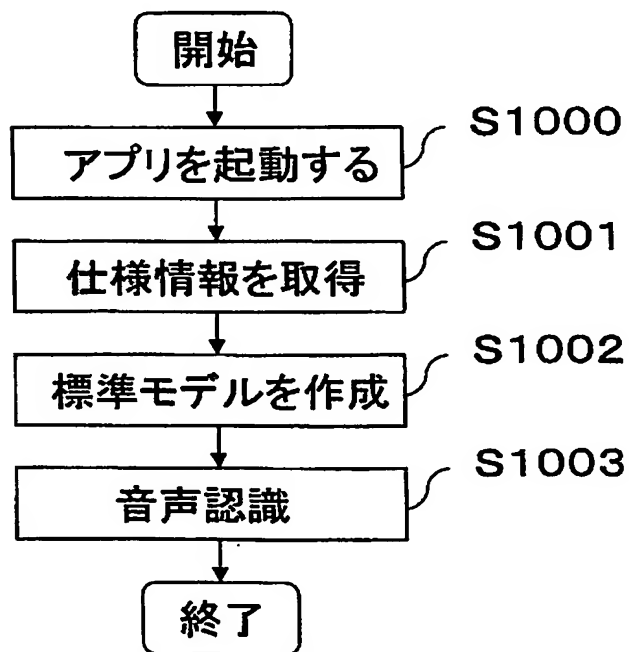
【図36】



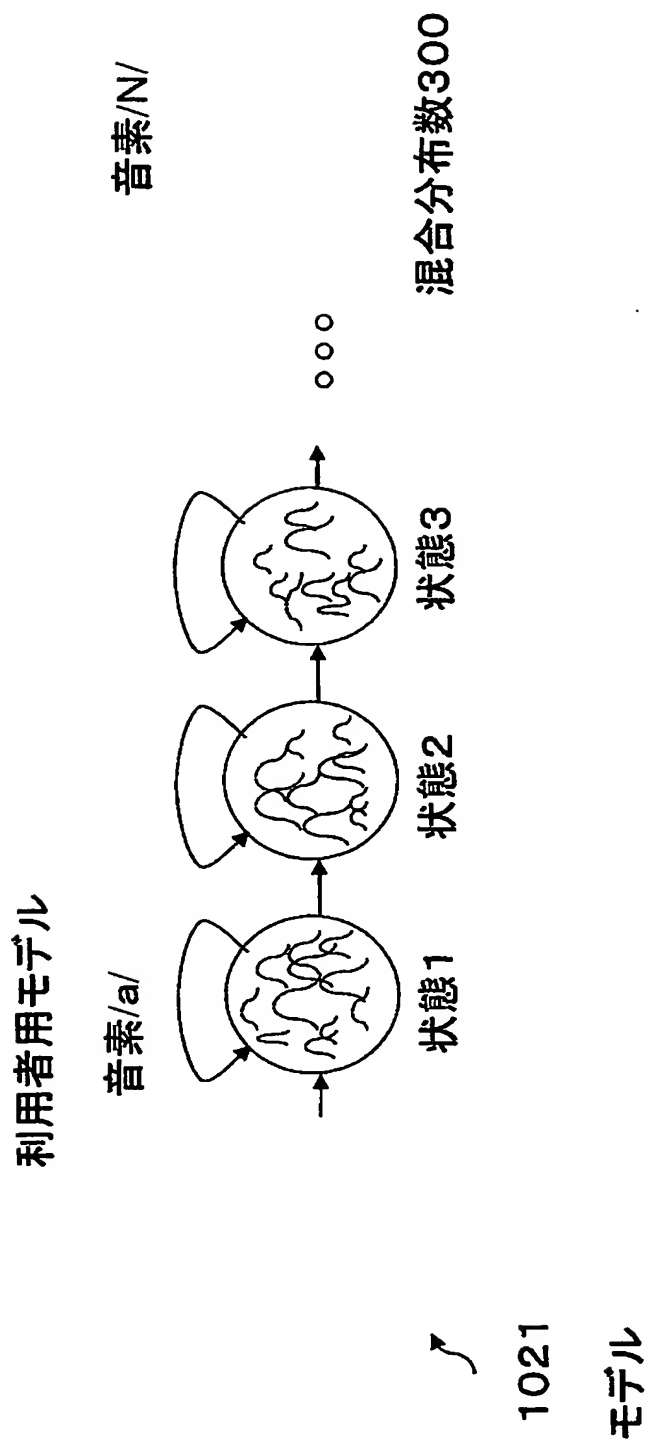
【図 37】

アプリケーション		仕様情報
ID	名前	
1	ゲームA	混合分布数3
2	ゲームB	混合分布数5
3	株取引	混合分布数126
4	テレビのリモコン	混合分布数5
5	翻訳	混合分布数64

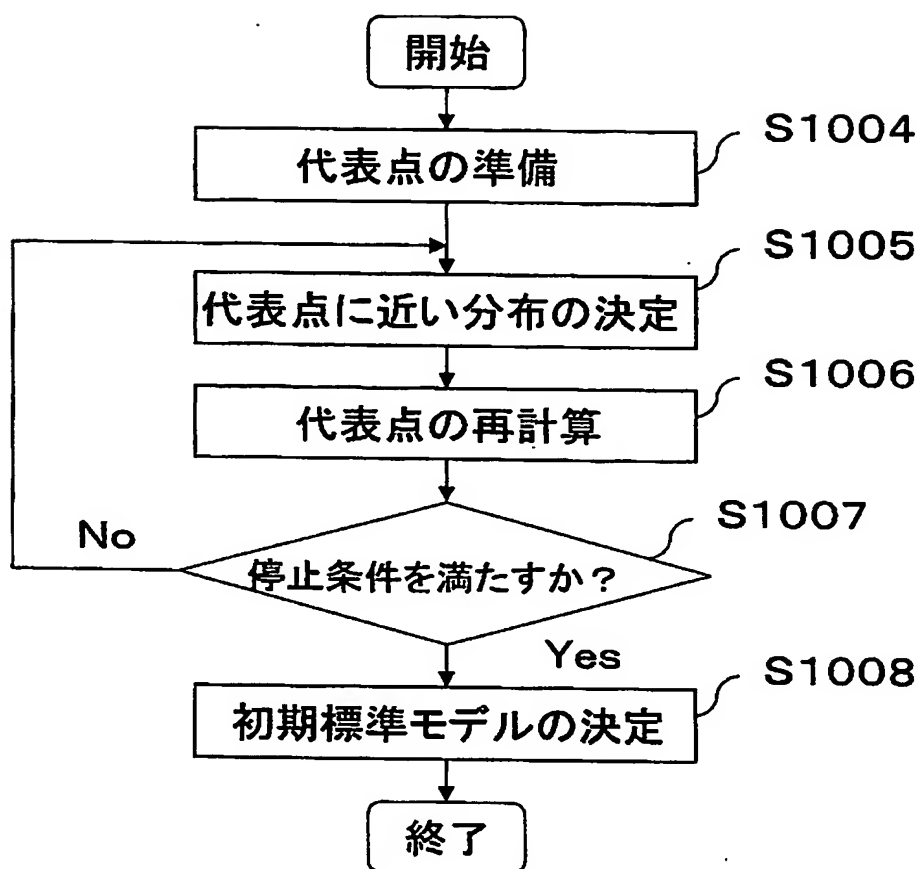
【図 38】



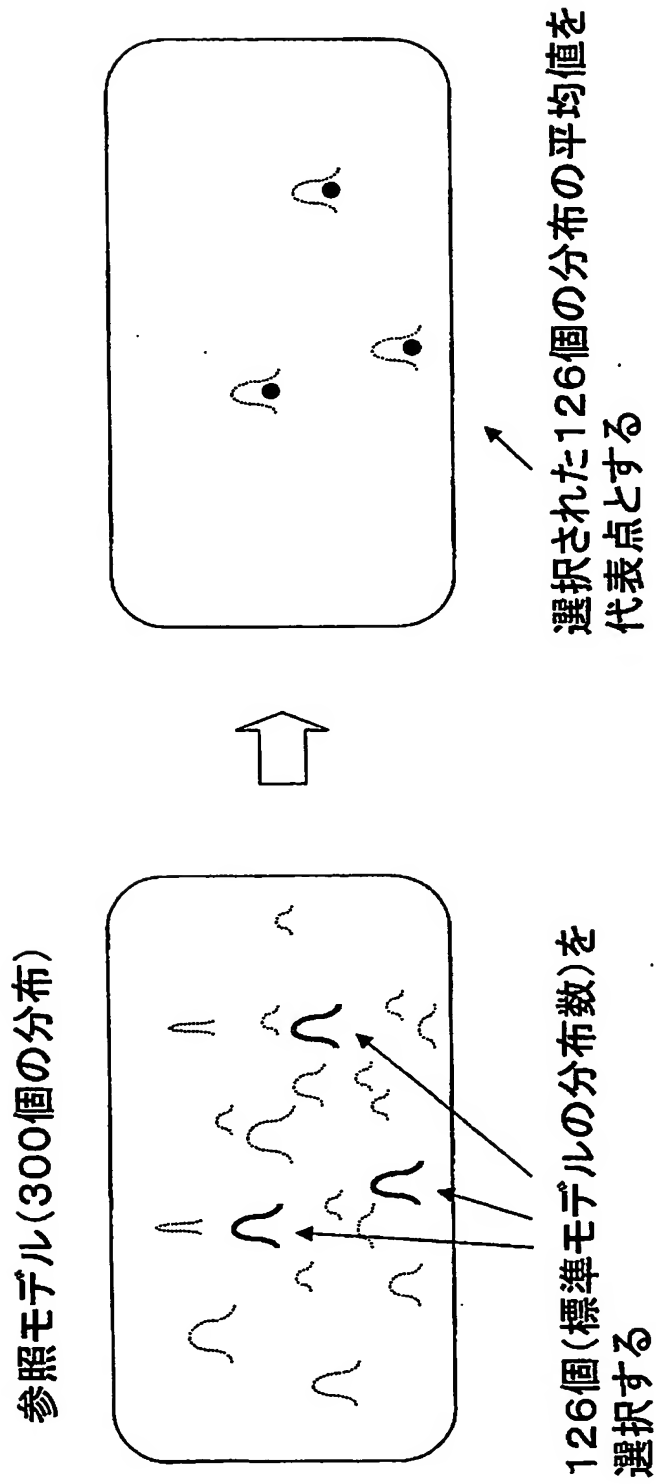
【図 39】



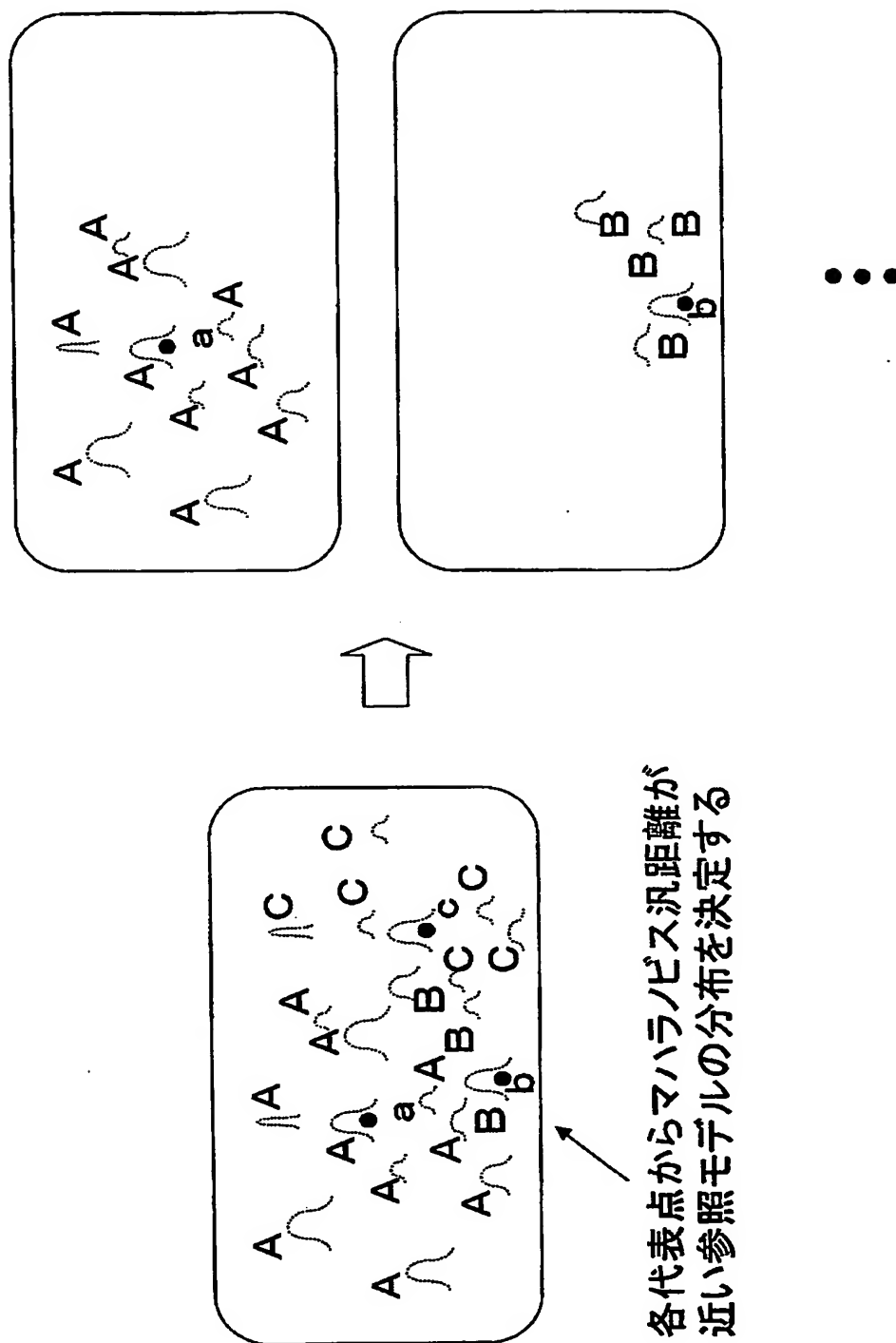
【図 40】



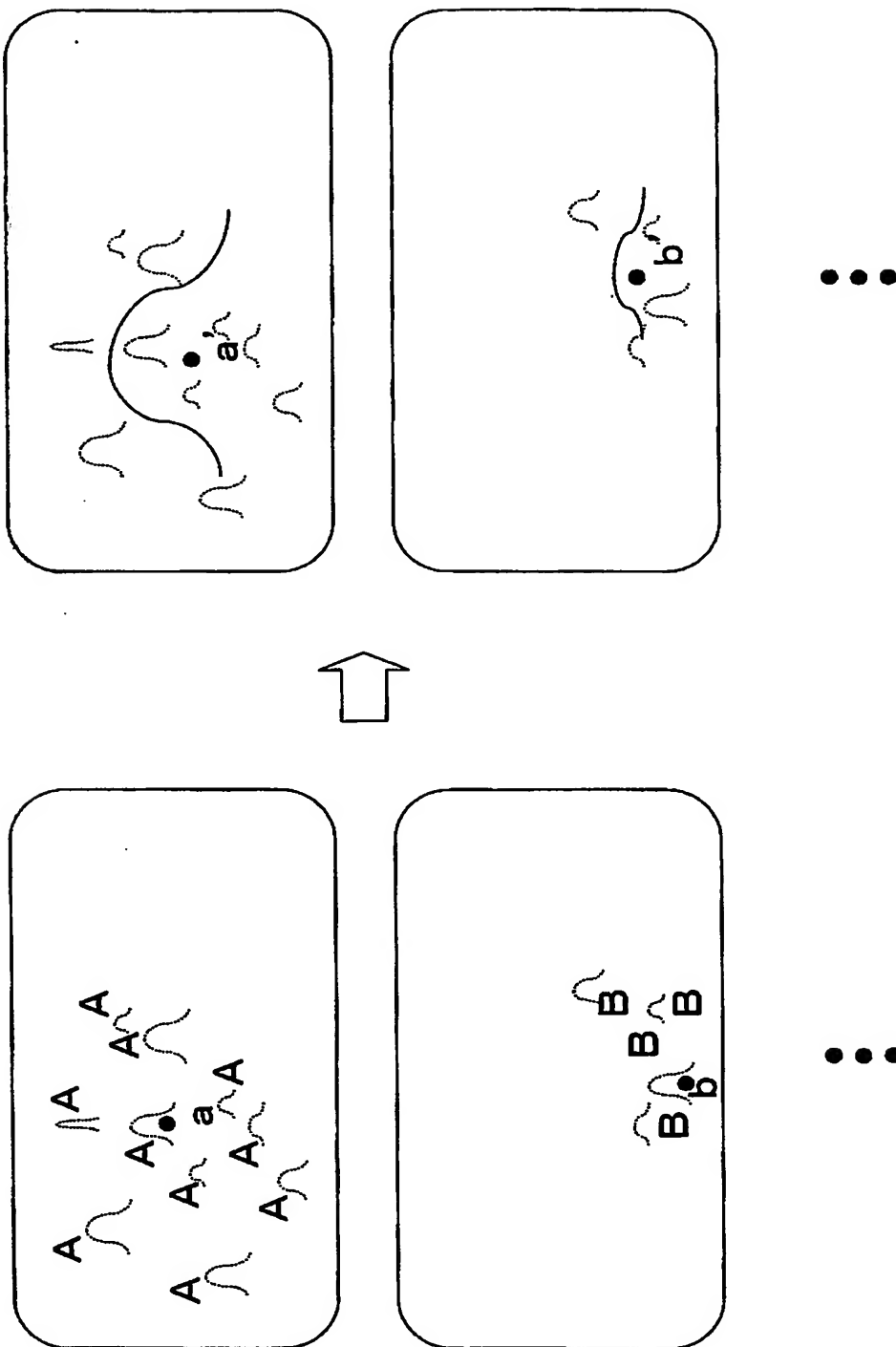
【図 41】



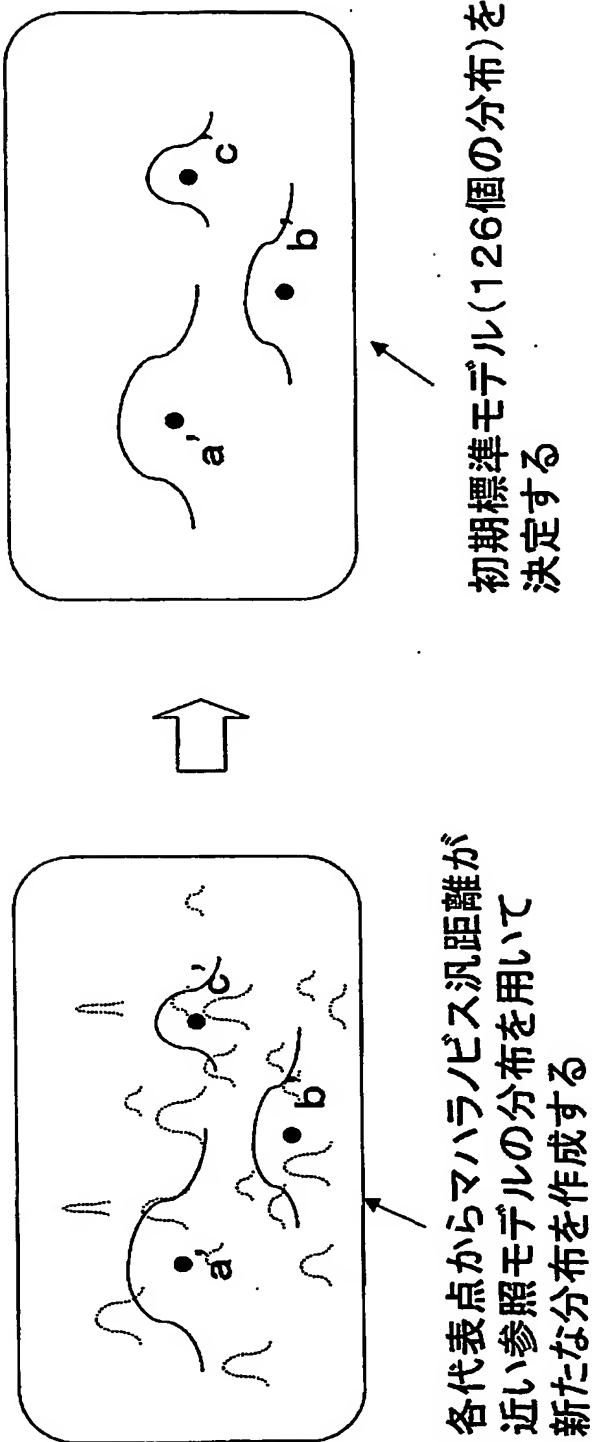
【図 4 2】



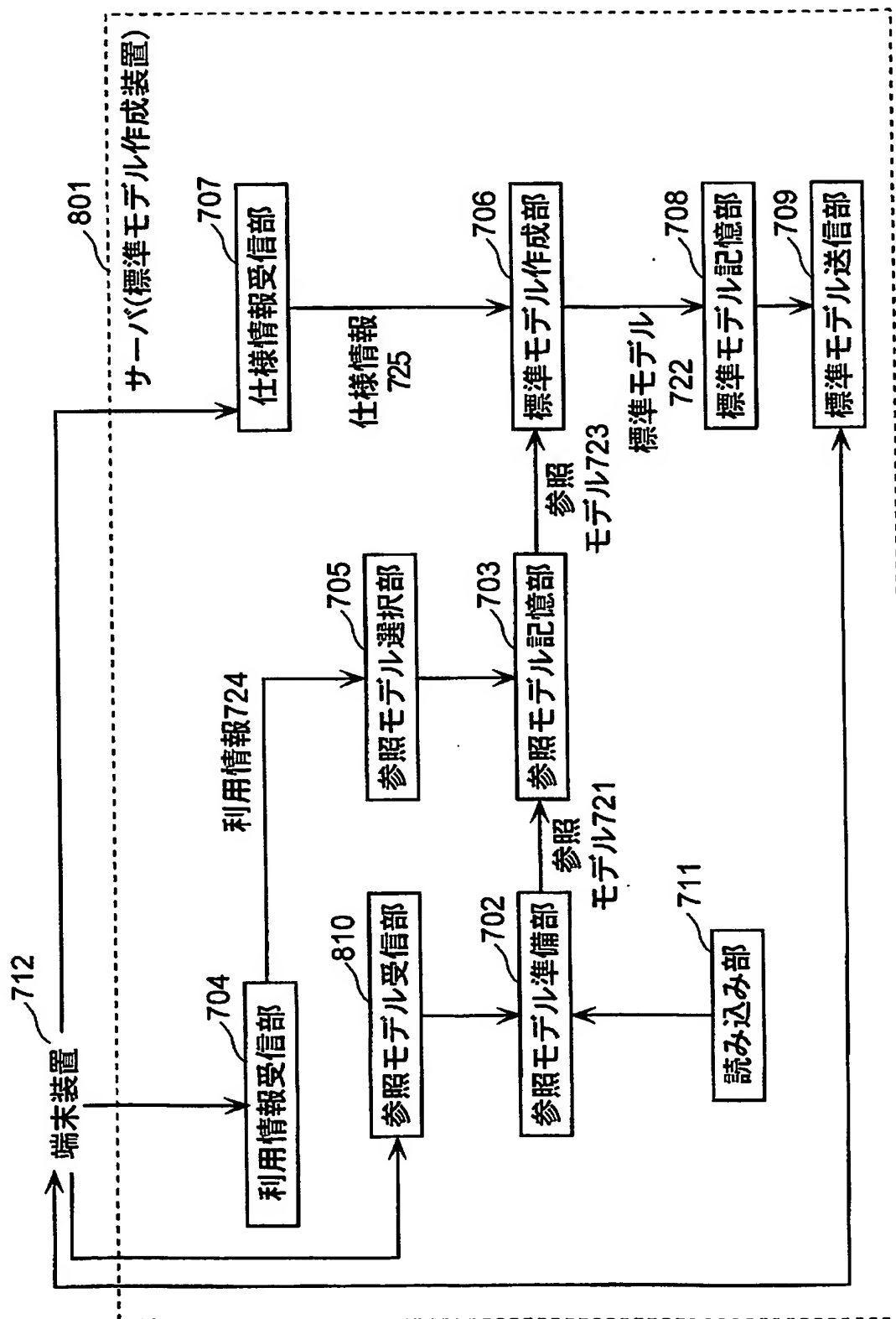
【図 43】



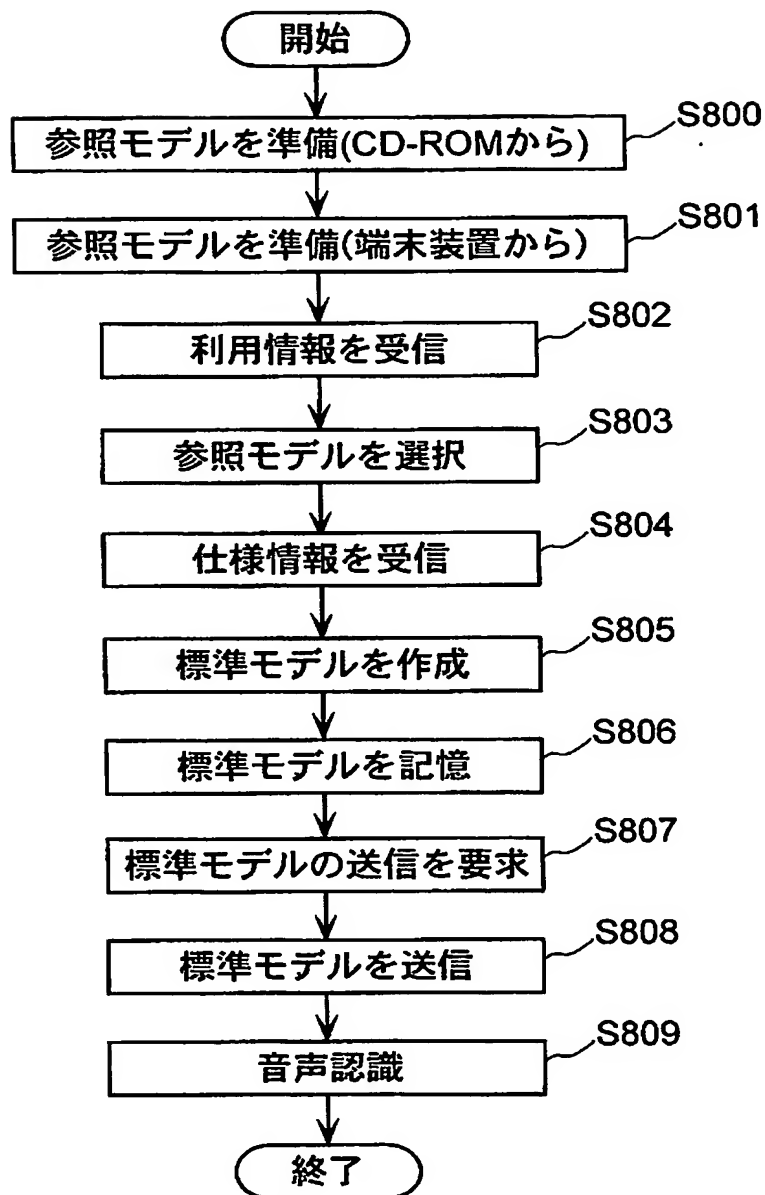
【図 44】



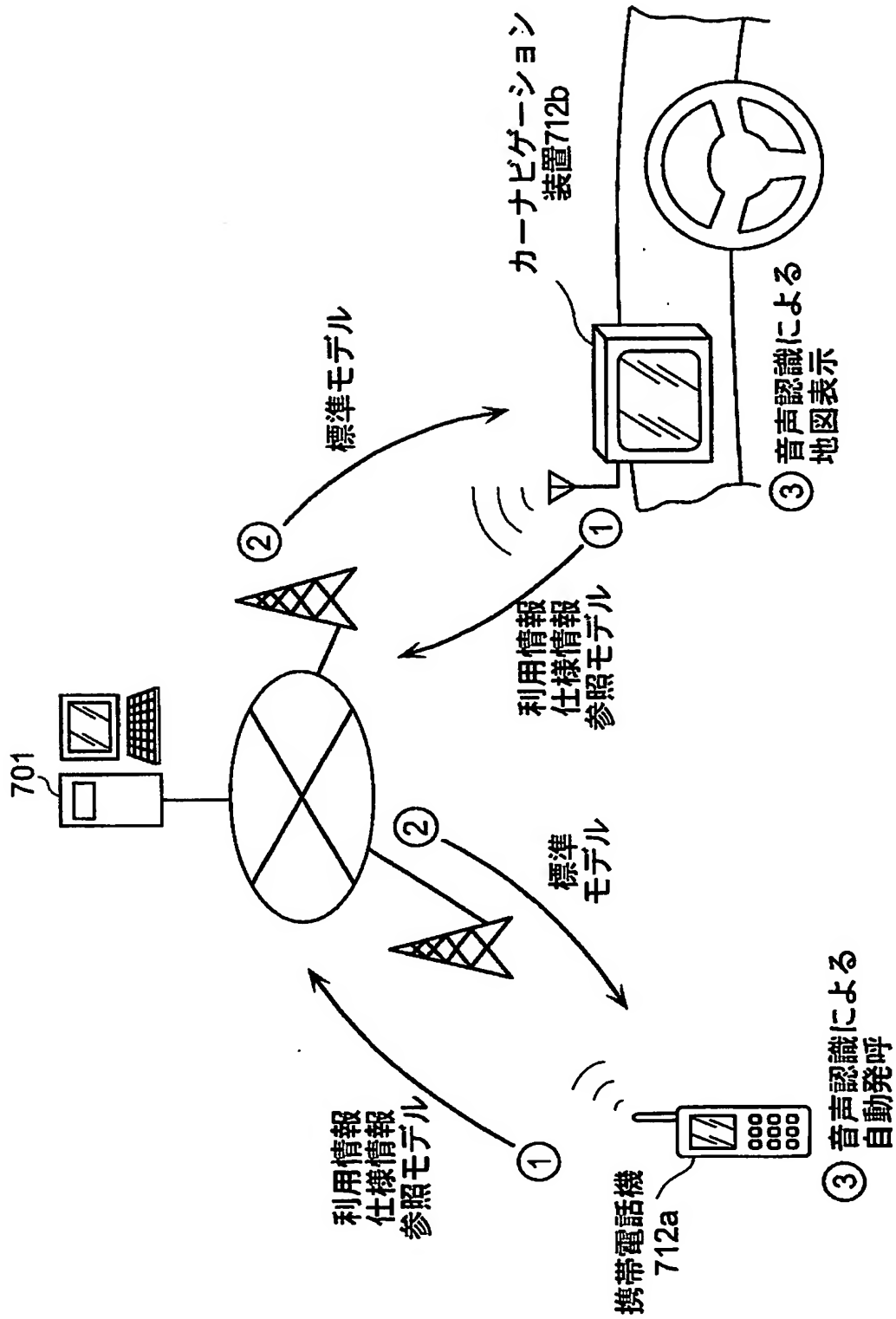
【図 45】



【図 4 6】



【図 47】

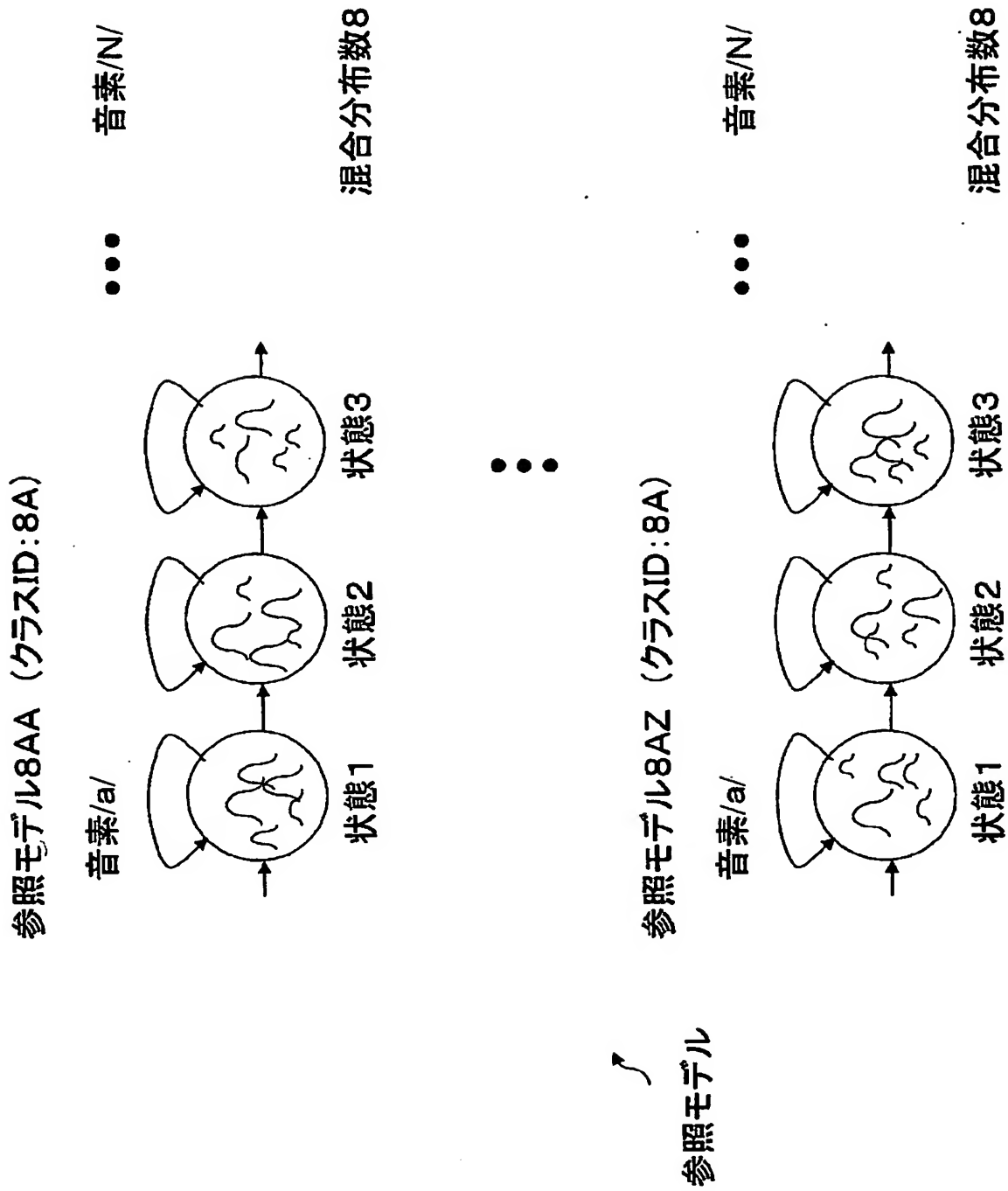


【図 48】

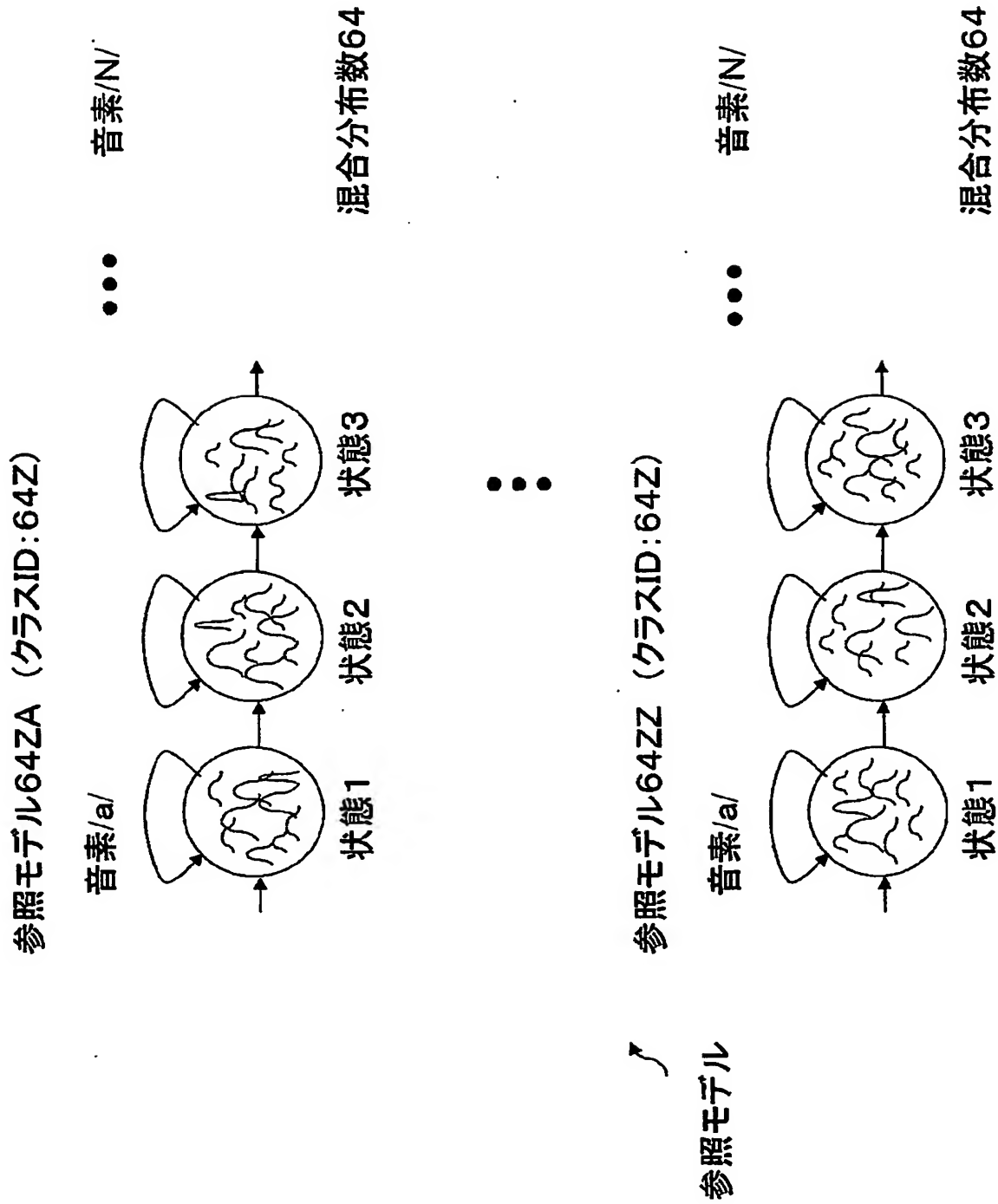
クラスID・初期標準モデル・参照モデル対応表

クラスID	初期標準モデル	参照モデル
8A	初期標準モデル8A	参照モデル8AA
		参照モデル8AB
		参照モデル8AC
		⋮
		参照モデル8AZ
⋮	⋮	⋮
64Z	初期標準モデル64Z	参照モデル64ZA
		参照モデル64ZB
		参照モデル64ZC
		⋮
		参照モデル64ZZ

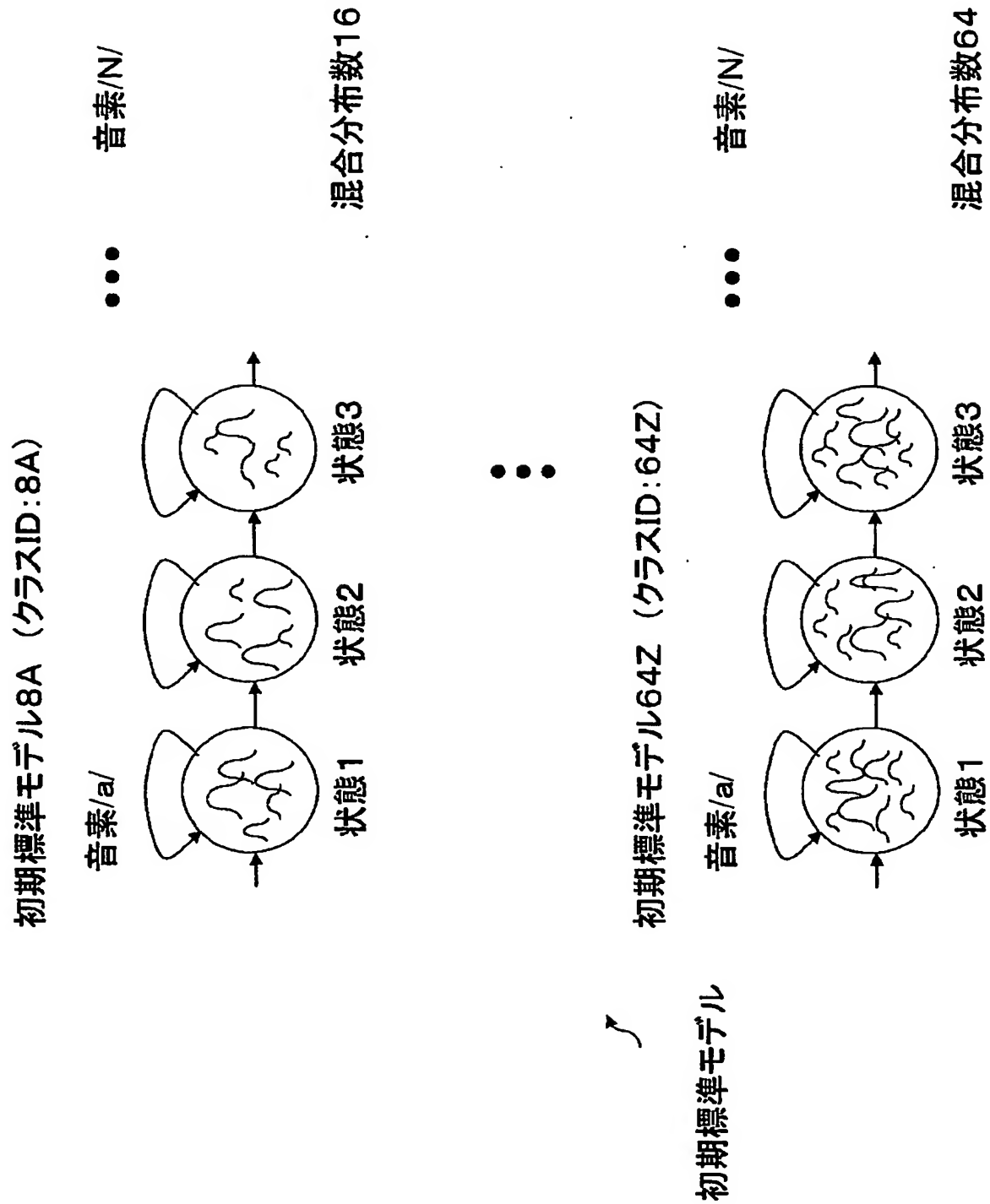
【図 49】



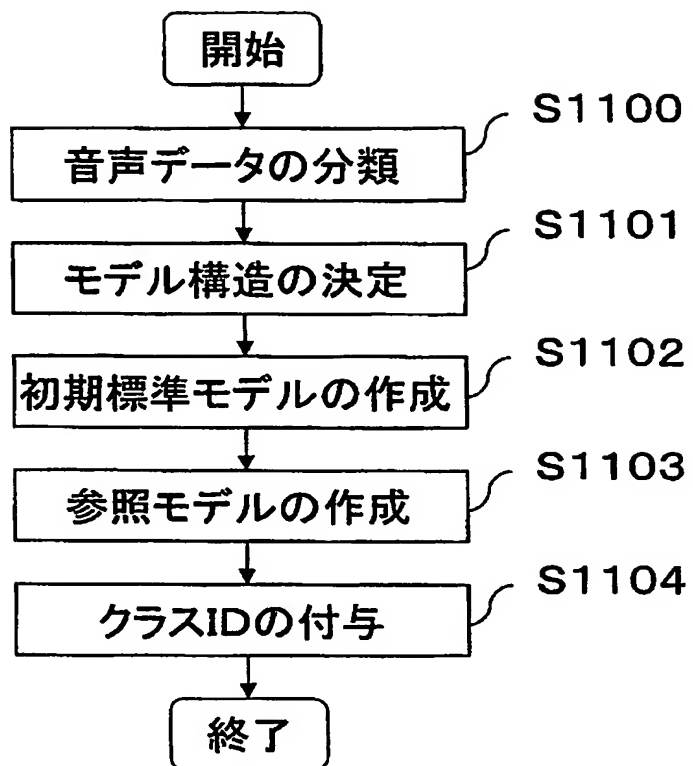
【図 50】



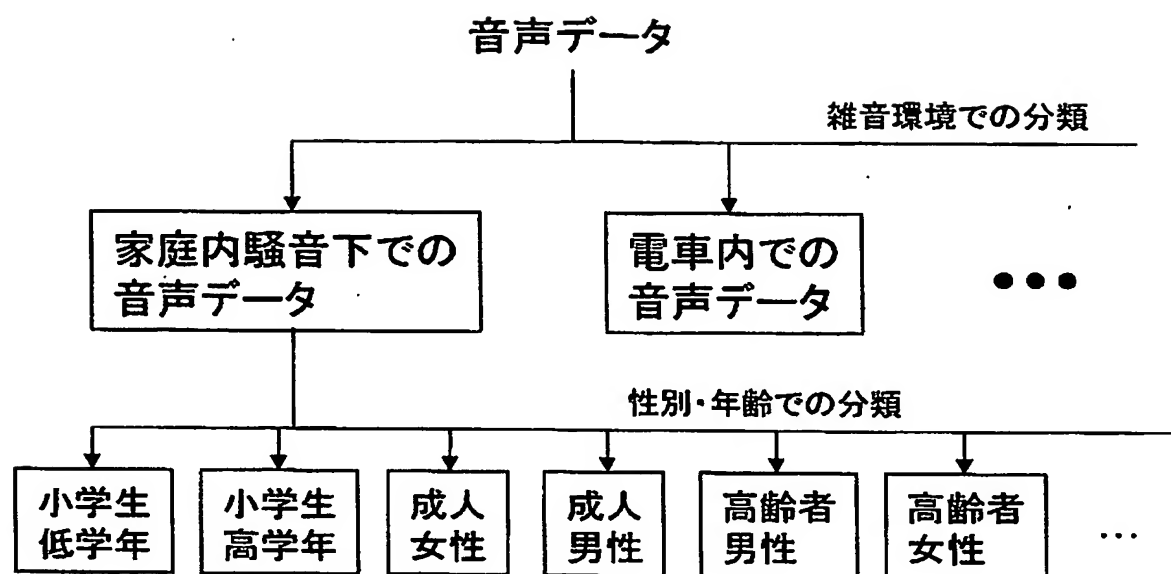
【図 51】



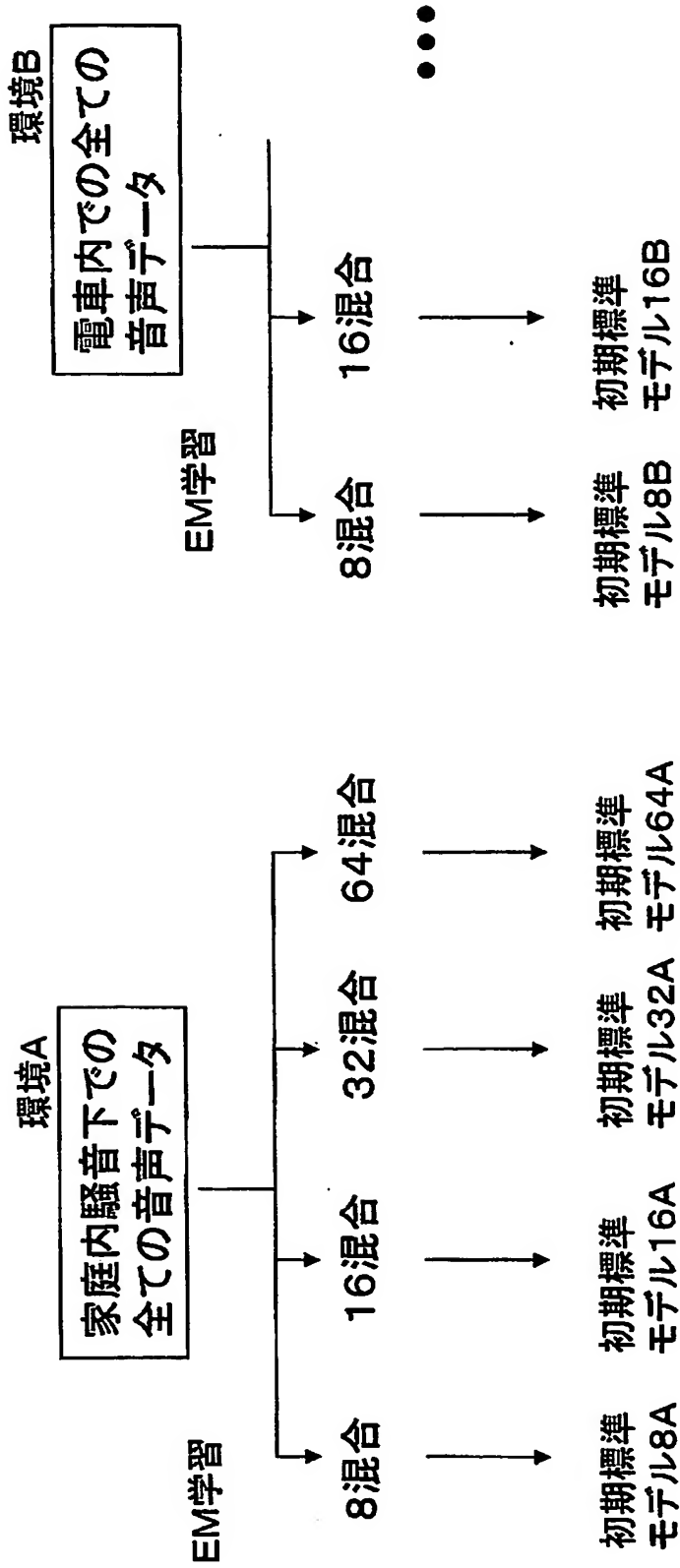
【図 5 2】



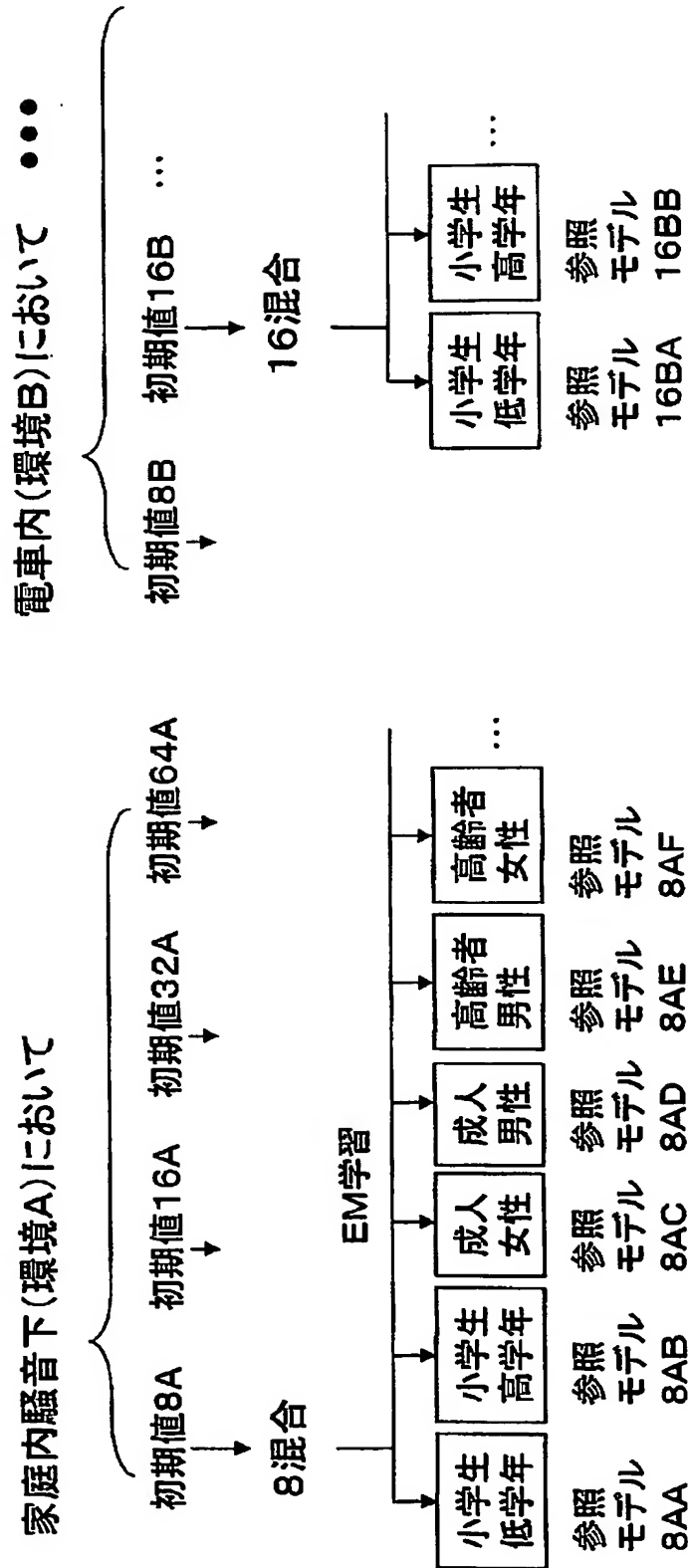
【図 53】



【図 54】



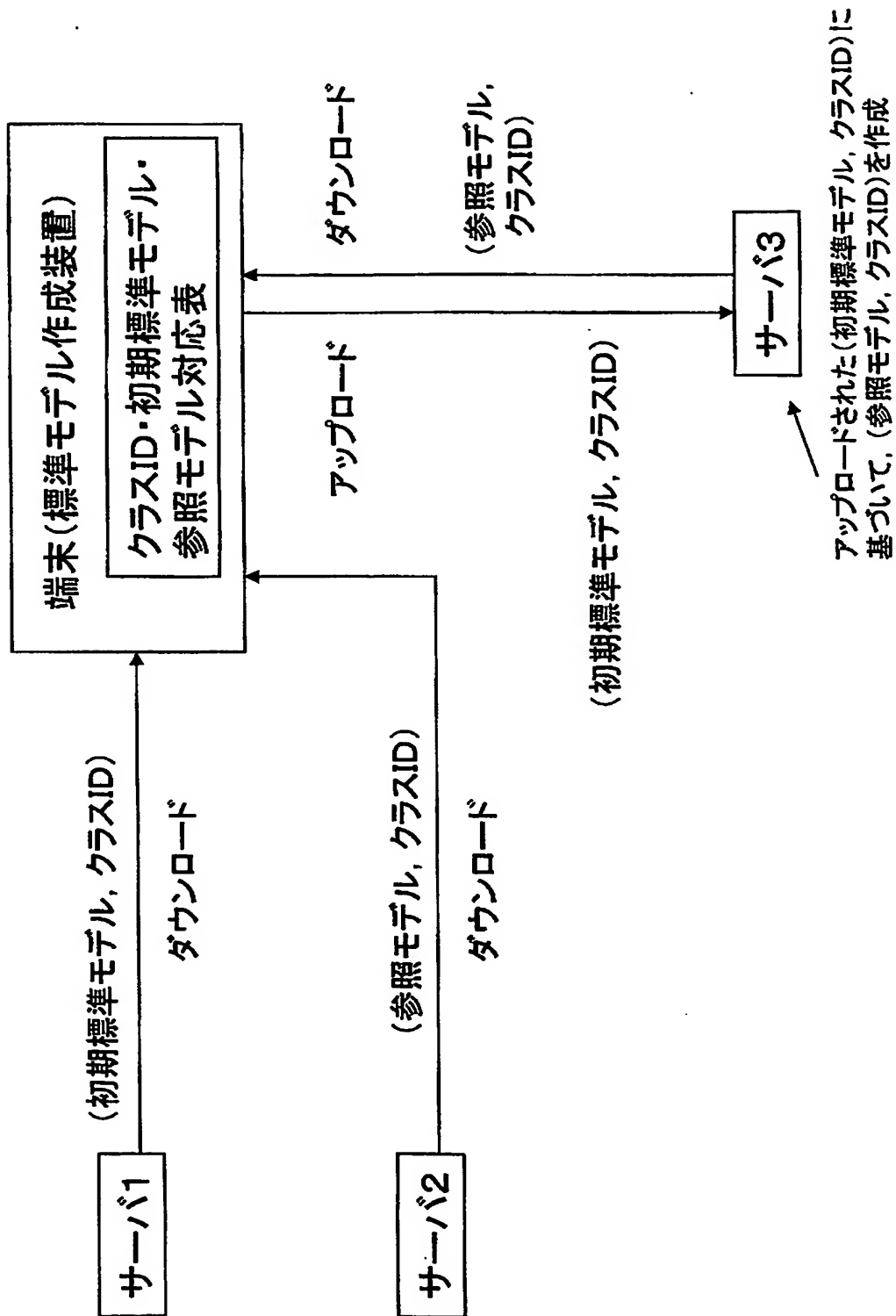
【図 55】



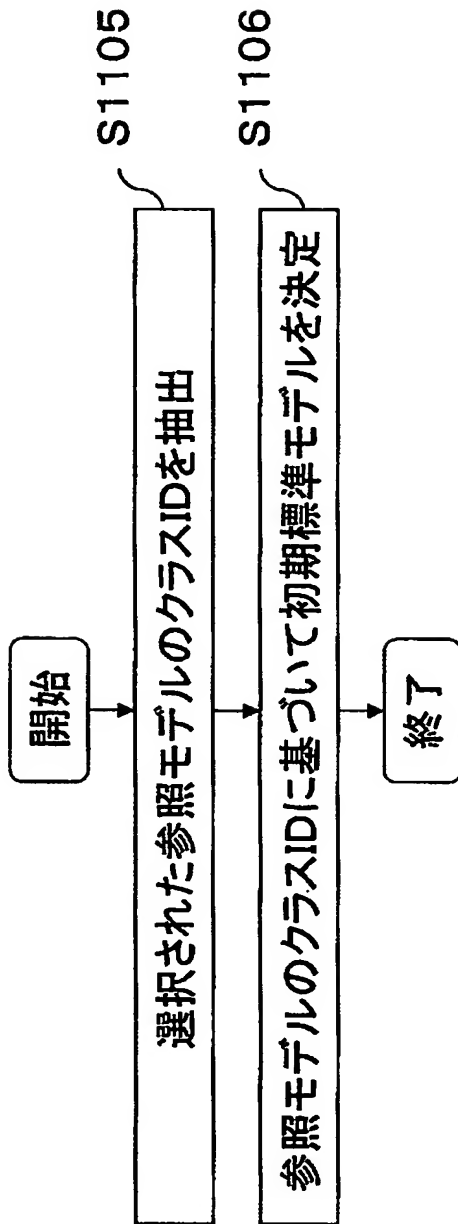
【図 56】

クラスID	初期標準モデル	参照モデル	備考 (参照モデルの特長)
8A	初期標準モデル8A	参照モデル8AA 参照モデル8AB 参照モデル8AC ⋮	家庭内騒音, 8混合, 小学校低学年 家庭内騒音, 8混合, 小学校高学年 家庭内騒音, 8混合, 成人女性 ⋮
16A	初期標準モデル16A	参照モデル16AA 参照モデル16AB 参照モデル16AC ⋮	家庭内騒音, 16混合, 小学校低学年 家庭内騒音, 16混合, 小学校高学年 家庭内騒音, 16混合, 成人女性 ⋮
⋮	⋮	⋮	⋮
64B	初期標準モデル64B	参照モデル64BA 参照モデル64BB 参照モデル64BC ⋮	電車内, 64混合, 小学校低学年 電車内, 64混合, 小学校高学年 電車内, 64混合, 成人女性 ⋮

【図 57】



【図 58】

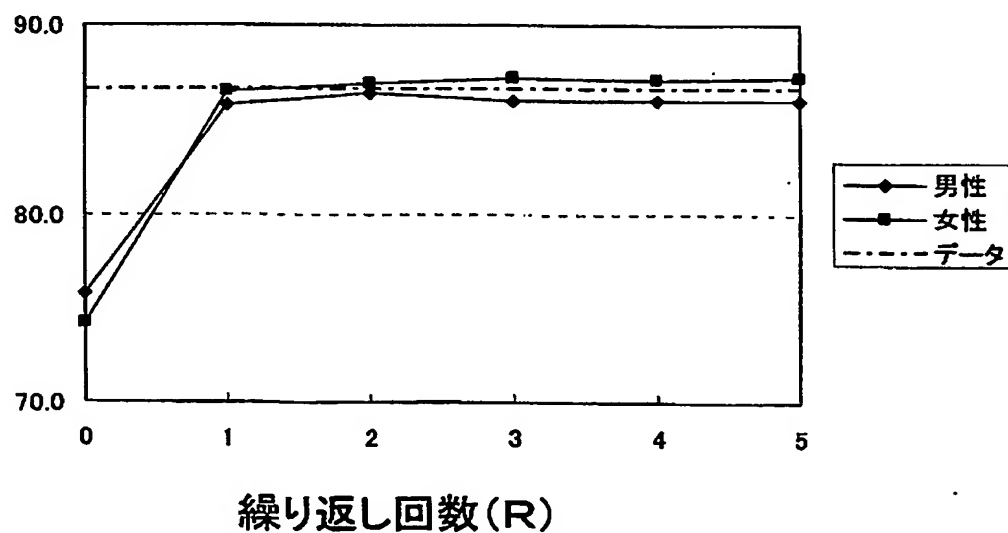


【図 59】

選択された参照モデル	クラスID
参照モデル8AA	8A
参照モデル16AA	16A
参照モデル16AB	16A
参照モデル16AC	16A
参照モデル16BA	16B
参照モデル64BA	64B

【図 60】

認識率(%)





【書類名】 要約書

【要約】

【課題】 隠れマルコフモデル、ベイズ理論、線形判別分析などの確率モデルによる音声認識、文字認識、画像認識などのパターン認識、ベイジアンネットなどの確率モデルによる意図理解、確率モデルによるデータマイニングなどに用いる高精度な標準モデルを提供する。

【解決手段】 1以上の参照モデルを準備する参照モデル準備部102と、参照モデル準備部102が準備した参照モデル121を記憶する参照モデル記憶部103と、参照モデル記憶部103が記憶している1以上の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデル122を作成する標準モデル作成部104とを備える。

【選択図】 図1



認定・付加情報

特許出願の番号	特願 2003-089179
受付番号	50300509093
書類名	特許願
担当官	第八担当上席 0097
作成日	平成15年 4月 4日

<認定情報・付加情報>

【提出日】 平成15年 3月27日

特願 2003-089179

出願人履歷情報

識別番号

[000005821]

1. 変更年月日

1990年 8月28日

[変更理由]

新規登録

住 所

大阪府門真市大字門真1006番地

氏 名

松下電器産業株式会社